

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

SUMMARIES

Run on: July 20, 2003, 05:49:35 : Search time 1537 Seconds
(without alignments)
8994.036 Million cell updates/sec

Title: US-09-719-002-1
Perfect score: 475
Sequence: 1 gaatcttattggacactga..... acaacacacacaatcatg 475

Scoring table: ORIGO.NUC
Gapop 60.0 , Gapext 60.0

Searched: 2054640 seqs, 14551402878 residues

Word size : 0

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : GenBank:
 1: gb_ba: *
 2: gb_htg: *
 3: gb_in: *
 4: gb_cm: *
 5: gb_ov: *
 6: gb_pat: *
 7: gb_ph: *
 8: gb_pl: *
 9: gb_pr: *
 10: qb_ro: *
 11: qb_sts: *
 12: qb_sy: *
 13: qb_un: *
 14: qb_vii: *
 15: em_ba: *
 16: em_fun: *
 17: em_hum: *
 18: em_in: *
 19: em_inu: *
 20: em_on: *
 21: em_or: *
 22: em_ov: *
 23: em_pat: *
 24: em_ph: *
 25: em_pl: *
 26: em_ro: *
 27: em_sts: *
 28: em_yi: *
 29: em_un: *
 30: em_htg_hum: *
 31: em_htg_inv: *
 32: em_htg_other: *
 33: em_htg_mus: *
 34: em_htg_pln: *
 35: em_htg_rdn: *
 36: em_htg_mam: *
 37: em_htg_vrt: *
 38: em_sy: *
 39: em_htg_hum: *
 40: em_htg_mus: *
 41: em_htg_other: *

Pred. No. is the number of results predicted by chance to have a

Result No.	Score	Query Match	Length	DB ID	Description
1	400	84.2	473	6 AX008564	AX008564 sequence
2	400	54.1	42	6 AX008582	AX008582 sequence
3	22	4.6	29	6 AX008576	AX008576 sequence
4	22	4.6	33	6 AX008574	AX008574 sequence
5	22	4.6	40	6 AX008579	AX008579 sequence
6	21	4.4	65361	2 AC011863	AC011863 Homo sapi
7	21	4.4	73218	2 AC111724	AC111724 Rattus no
8	21	4.4	92250	9 AC056693	AC056693 Homo sapi
9	21	4.4	122597	2 CNS05C8E	A1731874 Oriza sat
10	21	4.4	130732	2 AP03848	AP03848 Oriza sat
11	21	4.4	139613	2 CNS05C7T	AL131744 Oriza sat
12	21	4.4	149988	2 AC011929	AC011929 Homo sapi
13	21	4.4	156248	9 AC073322	AC073322 Homo sapi
14	21	4.4	162227	2 AC093224	AC093224 Rattus no
15	20	4.2	29	6 AX008577	AX008577 sequence
16	20	4.2	29	6 AX008578	AX008578 sequence
17	20	4.2	41	6 AX008575	AX008575 sequence
18	20	4.2	918	4 AF271901	AF271901 Bos tauru
19	20	4.2	7595	4 BTCAIR35	X11908 Bovine gene
20	20	4.2	39416	3 CEB0011	Z50634 Caenorhabdi
21	20	4.2	78448	2 AC111856	AC114856 Rattus no
22	20	4.2	99135	2 AC094561	AC094561 Homo sapi
23	20	4.2	111656	2 AC010449	AC010449 Homo sapi
24	20	4.2	128676	2 AC126877	AC126877 Rattus no
25	20	4.2	134955	10 AL714026	AL714026 Mouse DNA
26	20	4.2	135619	9 AL354809	AL354809 Human DNA
27	20	4.2	138996	4 AC008803	AC008803 Homo sapi
28	20	4.2	14841	2 AC121995	AC121995 Mus muscu
29	20	4.2	145658	9 AC093154	AC093154 Homo sapi
30	20	4.2	149751	2 AC043339	AC043339 Homo sapi
31	20	4.2	151111	2 AC103077	AC103077 Rattus no
32	20	4.2	157348	2 AF002446	AF002446 Homo sapi
33	20	4.2	162731	2 AC109162	AC109162 Mus muscu
34	20	4.2	154429	9 AC023366	AC023366 Homo sapi
35	20	4.2	166841	2 AC113523	AC113523 Rattus no
36	20	4.2	166960	2 AC015958	AC015958 Homo sapi
37	20	4.2	172189	9 AC105518	AC105518 Homo sapi
38	20	4.2	176224	2 AF000998	AF000998 Homo sapi
39	20	4.2	176999	2 AC053518	AC053518 Homo sapi
40	20	4.2	177444	2 AC084338	AC084338 Mus muscu
41	20	4.2	184436	9 AP000119	AP000119 Homo sapi
42	20	4.2	185178	2 AP000337	AP000337 Homo sapi
43	20	4.2	187310	2 AC101870	AC101870 Mus muscu
44	20	4.2	192519	10 MMU178435	AC1278435 Mus muscu
45	20	4.2	195907	2 AC095185	AC095185 Rattus no

ALIGNMENTS

RESULT 1
AX008564 LOCUS AX008564
DEFINITION Sequence 1 from Patent WO9960507.
ACCESSION AX008564
VERSION AX008564.1 GI:9996114
KEYWORDS
ORGANISM garden asparagus
Asparagus officinalis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales;
Asparagaceae; Asparagis.
REFERENCE 1 (bases 1 to 473)
AUTHORS Draper,J., Kenton,P. and Paul,W.
TITLE Inducible promoters

JOURNAL Patent: WO 9966057 A 1 23-DEC-1999;
 DRAPER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL
 WYATT (GB) Location/Qualifiers

FEATURES source 1. .473
 /organism="Asparagus officinalis"
 /db_xref="taxon:4686" ORIGIN

Query Match 84.2%; Score 400; DB 6; Length 473;
 Best Local Similarity 100.0%; Pred. No. 3.5e-209;
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAATTCCTATTGGGACCCGACTCTCTTGTGTTGCTGCCGAGGTGCGTGSAAATTCTGT 60
 1 GAATTCCTATTGGGACCCGACTCTCTTGTGTTGCTGCCGAGGTGCGTGSAAATTCTGT 60
 Db 61 TGGCACACATCTGGCTTGCTGATGACGACTCCATATTTCAGTCAT 120
 61 TGGCACACATCTGGCTTGCTGATGACGACTCCATATTTCAGTCAT 120
 Qy 121 GAGAGAACCATGACTRAAGAATTGCTTACCCATAAACCTAACAGAGAT 180
 121 GAGAGAACCATGACTRAAGAATTGCTTACCCATAAACCTAACAGAGAT 180
 Qy 181 GAGACATCACAGAAATACTTAAATAGTCTTGGGTGAAATTGGAACATGATC 240
 181 GAGACATCACAGAAATACTTAAATAGTCTTGGGTGAAATTGGAACATGATC 240
 Db 241 CTGATTAATTAACCTTGTGAAATAAATTAAGAAGTCTTACAGAAGACTGT 300
 241 CTGATTAATTAACCTTGTGAAATAAATTAAGAAGTCTTACAGAAGACTGT 300
 Qy 301 TCTTACATGAGACTAGTCAGCAAGCTTACCTTATTCGCAAGGCTTACAGTCA 360
 301 TCTTACATGAGACTAGTCAGCAAGCTTACCTTATTCGCAAGGCTTACAGTCA 360
 Db 361 CAATCGAGATTCCCATGGACTGATGGACACCATCAA 400
 361 CAATCGAGATTCCCATGGACTGATGGACACCATCAA 400
 Qy 361 CAATCGAGATTCCCATGGACTGATGGACACCATCAA 400
 Db 361 CAATCGAGATTCCCATGGACTGATGGACACCATCAA 400

RESULT 2
 AX008582/c
 LOCUS AX008582
 DEFINITION Sequence 19 from Patent WO9966057.
 ACCESSION AX008582.1
 VERSION GI:9996132
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 33)
 AUTHORS Draper,J., Kenton,P. and Paul,W.
 TITLE Inducible Promoters
 JOURNAL Patent: WO 9966057-A 1 23-DEC-1999;
 DRAPER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL
 WYATT (GB) Location/Qualifiers

FEATURES source 1. .42
 /organism="synthetic construct"
 /db_xref="taxon:32630"/note="primer" ORIGIN

Query Match 4.6%; Score 22; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.2e-002;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTCTPATGGGACCTGACTCT 25
 Db 12 TTCTPATGGGACCTGACTCT 33

Query Match 5.1%; Score 24; DB 6; Length 42;
 Best Local Similarity 100.0%; Pred. No. 0.092;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

BASE COUNT 13 a 7 c 14 g 8 t
 ORIGIN

Query Match 5.1%; Score 24; DB 6; Length 42;
 Best Local Similarity 100.0%; Pred. No. 0.092;
 Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 GTCCACGAGACTCTACTTATCC 340
 |||||||

Db 42 GTCCACGAGACTCTACTTATCC 19
 RESULT 3
 AX008576
 LOCUS AX008576
 DEFINITION Sequence 13 from Patent WO9966057.
 ACCESSION AX008576
 VERSION AX008576.1 GI:9996126
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 29)
 AUTHORS draper,J., Kenton,P. and Paul,W.
 TITLE Inducible Promoters
 JOURNAL Patent: WO 9966057-A 13 23-DEC-1999;
 DRAPER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL
 WYATT (GB) Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"/note="primer" ORIGIN

Query Match 4.5%; Score 22; DB 6; Length 29;
 Best Local Similarity 100.0%; Pred. No. 1.2e-002;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 110 TTCCATGTCATGAGAAGCAC 131
 Db 8 TTCCATGTCATGAGAAGCAC 29

RESULT 4
 AX008574
 LOCUS AX008574
 DEFINITION Sequence 11 from Patent WO9966057.
 ACCESSION AX008574
 VERSION AX008574.1 GI:9996124
 KEYWORDS synthetic construct.
 SOURCE synthetic construct.
 ORGANISM synthetic construct.
 REFERENCE 1 (bases 1 to 33)
 AUTHORS Draper,J., Kenton,P. and Paul,W.
 TITLE Inducible Promoters
 JOURNAL Patent: WO 9966057-A 11 23-DEC-1999;
 DRAPER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL
 WYATT (GB) Location/Qualifiers

FEATURES source 1. .33
 /organism="synthetic construct"
 /db_xref="taxon:32630"/note="primer" ORIGIN

Query Match 4.6%; Score 22; DB 6; Length 33;
 Best Local Similarity 100.0%; Pred. No. 1.2e-002;
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 TTCTPATGGGACCTGACTCT 25
 Db 12 TTCTPATGGGACCTGACTCT 33

RESULT 5
 AX008579/c
 LOCUS AX008579
 DEFINITION Sequence 16 from Patent WO9966057.

ACCESSION AX008579
 VERSION AX008579.1 GI:9996129
 KEYWORDS synthetic construct
 SOURCE synthetic construct
 ORGANISM artificial sequences.
 REFERENCE 1 (bases 1 to 40)
 AUTHORS Draper,J., Kenton,P. and Paul,W.
 TITLE Inducible promoters
 JOURNAL Patent: WO 996057-A 16-23-DEC-1999;
 DRAPER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL
 WIATT (GB)
 FEATURES location/Qualifiers
 source 1..40
 /organism="synthetic construct"
 /db_xref="Taxon:32630"
 /notes="primer"
 BASE COUNT 4 a 6 c 14 g 16 t
 ORIGIN
 Query Match 4.6%; Score 22; DB 6; Length 40;
 Best Local Similarity 100.0%; Pred. No. 1.2; Matches 0; Indels 0; Gaps 0;
 Matches 22; Conservative 0; Mismatches 0;
 QY 451 CAAAAACACACACACACCACTC 472
 Db 40 CAAAAACACACACACACCACTC 19
 RESULT 6
 AC091863
 LOCUS AC091863 65361 bp DNA linear HTG 09-JUN-2001
 DEFINITION Homo sapiens chromosome 5 clone CTD-235015, WORKING DRAFT SEQUENCE,
 14 unordered pieces.
 ACCESSION AC091863
 VERSION AC091863.1 GTR:14333799
 AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 TITLE Sequencing of Human Chromosome 5
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 65361)
 AUTHORS DOE Joint Genome Institute.
 TITLE Direct Submission
 JOURNAL Submitted (09-JUN-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
 COMMENT -----Genome Center: Center: Joint Genome Institute
 Center Code: JGI
 Web site: <http://www.Jgi.doe.gov>
 Project Information
 Center Project Name: 762146
 Center clone name: CITB_HL_235015

 Summary Statistics
 Consensus quality: 54296 bases at least 940
 Consensus quality: 56671 bases at least 930
 Consensus quality: 57432 bases at least 920
 Consensus quality: 57432 bases at least 920
 Estimated insert size: 168010; agarose-fp estimation
 Estimated insert size: 64051; sum-of-contigs estimation
 Quality coverage: 5.38 in Q20 bases; agarose-fp estimation
 Quality coverage: 14.12 in Q20 bases; sum-of-contigs estimation.
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 14 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.

as soon as it is available and the accession number will
 be preserved.
 * 1 1343: contig of 1343 bp in length
 * 1 1344 1443: gap of unknown length
 * 1 1444 2340: contig of 1097 bp in length
 * 1 2541 2640: gap of unknown length
 * 1 2641 3711: contig of 1071 bp in length
 * 1 3712 3811: gap of unknown length
 * 1 3812 4821: contig of 1010 bp in length
 * 1 4822 4921: gap of unknown length
 * 1 4922 6127: contig of 1206 bp in length
 * 1 6128 6228 6227: gap of unknown length
 * 1 11911 11910: gap of 5683 bp in length
 * 1 12011 12010: gap of unknown length
 * 1 1447: contig of 2837 bp in length
 * 1 14848 1476: gap of unknown length
 * 1 14948 18476: contig of 3529 bp in length
 * 1 18477 18576: gap of unknown length
 * 1 18577 25093: contig of 6517 bp in length
 * 1 23094 25193: gap of unknown length
 * 1 25194 33110: contig of 7917 bp in length
 * 1 33111 33110: gap of unknown length
 * 1 32211 39795: contig of 6585 bp in length
 * 1 39795 39895: gap of unknown length
 * 1 39896 46442: contig of 6847 bp in length
 * 1 46442 46442: gap of unknown length
 * 1 46743 47022: contig of 1090 bp in length
 * 1 47022 57222: gap of unknown length
 * 1 57222 57923 57923: gap of 57923 bp in length
 * 1 57923 57923: contig of 7439 bp in length.
 RESULT 7
 AC121724
 LOCUS AC121724 73218 bp DNA linear HTG 24-JUL-2002
 DEFINITION Rattus norvegicus clone CH230-332D19, *** SEQUENCING IN PROGRESS
 ***, 39 unordered pieces.
 ACCESSION AC121724
 VERSION AC121724.2 GTR:23909372
 FEATURES location/Qualifiers
 source 1..65361
 /organism="Homo sapiens"
 /db_xref="Taxon:9806"
 /chromosome="5"
 /clone="CTD-235015"
 /clone.lib="CalTech human BAC library D"
 BASE COUNT 19445 a 13085 c 11801 g 19653 t 1377 others
 ORIGIN
 Query Match 4.4%; Score 21; DB 2; Length 65361;
 Best Local Similarity 100.0%; Pred. No. 4.1; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 241 CTACATTAATACTACACTTTC 261
 Db 16620 CTACATTAATACTACACTTTC 16640
 RESULT 7
 AC121724
 LOCUS AC121724 73218 bp DNA linear HTG 24-JUL-2002
 DEFINITION Rattus norvegicus clone CH230-332D19, *** SEQUENCING IN PROGRESS
 ***, 39 unordered pieces.
 ACCESSION AC121724
 VERSION AC121724.2 GTR:23909372
 FEATURES location/Qualifiers
 source 1..65361
 /organism="Homo sapiens"
 /db_xref="Taxon:9806"
 /chromosome="5"
 /clone="CTD-235015"
 /clone.lib="CalTech human BAC library D"
 BASE COUNT 19445 a 13085 c 11801 g 19653 t 1377 others
 ORIGIN
 Query Match 4.4%; Score 21; DB 2; Length 65361;
 Best Local Similarity 100.0%; Pred. No. 4.1; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 241 CTACATTAATACTACACTTTC 261
 Db 16620 CTACATTAATACTACACTTTC 16640
 RESULT 7
 AC121724
 LOCUS AC121724 73218 bp DNA linear HTG 24-JUL-2002
 DEFINITION Rattus norvegicus
 SOURCE Rattus norvegicus
 ORGANISM Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus
 REFERENCE 1 (bases 1 to 73218)
 AUTHORS Munro,D.M., Adams,C., Adio-Oduola,B., Ali-Osman,F.R., Allen,C.,
 Alsbrooks,S.L., Amaral,R.C., Areo-J.R., Arele,M., Banks,T.,
 Barbara,J., Benton,J., Blinige,K., Blankenburg,K., Bonin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buhay,C., Burck,P., Burkett,L.K., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,G.,
 Cleveland,C.D., Cox,M.D., Dathorne,S.R., David,R.,
 Davis,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,R., Duran-Rocha,S., Durbin,K.J.,

FEATURES * 68950 73218: contig of 4269 bp in length.
 FEATURES source
 source 1. .73218
 /organism="Rattus norvegicus"
 /clone="CH230-332B19"
 /db_xref="taxon:10116"
 BASE COUNT 20948 a 15112 c 14120 g 19150 t 3888 others
 ORIGIN

Query Match 4.4%; Score 21; DB 2; Length 73218;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 250 TTACAACTTTGCATAAA 270
 |||||||
 Db 27829 TTACAACTTTGCATAAA 27849

RESULT 8
 AC026693
 LCUS
 DEFINITION Homo sapiens chromosome 5 Clone CTC-261E10, complete sequence.
 ACCESSION AC026693
 VERSTON 45 GI:19887318
 KEYWORDS HTG,
 SOURCE
 ORGANISM Homo sapiens.
 REFERENCE
 AUTHORS Bukayoya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL Direct Submission
 REFERENCE 1 (bases 1 to 92250)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 92250)
 AUTHORS DOE Joint Genome Institute.
 JOURNAL Direct Submission
 REFERENCE 3 (bases 1 to 92250)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL Direct Submission
 REFERENCE 4 (bases 1 to 92250)
 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.
 JOURNAL Direct Submission
 COMMENT On Oct 3, 2001 this sequence version replaced gi:10044353.
 Draft Sequence produced by DOE Joint Genome Institute
 www.shgc.stanford.edu
 Finishing completed at Stanford Human Genome Center
 Quality: Phrap Quality >=40 99.5% of Sequence;
 Estimated Total Number of Errors is 0.4.
 FEATURES source
 source 1. .92250
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="5"
 /clone="CTC-261E10"
 BASE COUNT 28028 a 17238 c 17152 g 29832 t
 ORIGIN

Query Match 4.4%; Score 21; DB 9; Length 92250;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 241 CTACATTAACTTACAACTTG 261
 |||||||
 Db 58285 CTACATTAACTTACAACTTG 58305

RESULT 9
 CNS08C8E/c
 LOCUS CNS08C8E
 DEFINITION Oryza sativa chromosome 12 clone OSJNBB0077C18, *** SEQUENCING IN PROGRESS ***,
 4 ordered pieces.
 ACCESSION AL731874
 VERSION AL731874.1 GI:2032824
 KEYWORDS HTG; HGVS_PPHASE2; HGVS_ACTIVEPIN.
 SOURCE
 ORGANISM Oryza sativa.
 Oryza sativa; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophytina; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartioideae; Oryzeae; Oryza.
 1 (bases 1 to 122697)
 Choisne,N., Orjeda,G., Cattolico,L., Damande,N., Wincker,P.,
 Seguret,B., Pelletier,E., Scarpetti,C., Salanoubat,M.,
 Weissenbach,J. and Quétier,F.
 Oryza sativa chromosome 12 sequencing
 TITLE Unpublished
 JOURNAL
 REFERENCE 2 (bases 1 to 122697)
 AUTHORS Genoscope.
 JOURNAL Direct Submission
 COMMENT Submitted (29-APR-2002) Genoscope - Centre National de Séquençage - FRANCE (E-mail : seqref@genoscope.cnrs.fr)
 - Web : www.genoscope.cnrs.fr
 IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence.
 Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage, etc.
 Contigs composition :
 11793 bp contig from 1 to 11793
 2451 bp contig from 1184 to 14344
 11789 bp contig from 14445 to 26233
 96364 bp contig from 26334 to 122697.
 * NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have provided by the submitter.
 * This sequence will be replaced by the finished sequence as soon as it is available and the accession number will be preserved.
 * 1 11793: contig of 11793 bp in length
 * 11794 11893: gap of 100 bp
 * 11894 14344: contig of 2451 bp in length
 * 14345 14444: gap of 100 bp
 * 14445 26233: contig of 11789 bp in length
 * 26234 26333: gap of 100 bp
 * 26334 122697: contig of 96364 bp in length.
 FEATURES source
 source 1. .122697
 /organism="Oryza sativa"
 /cultivar="Nipponbare"
 /sub_species="japonica"
 /db_xref="taxon:4530"
 /chromosome="12"
 /clone="OSJNBB0077C18"
 /clone_1_id="OSJNBB"
 BASE COUNT 35100 a 27094 c 26028 g 34171 t 304 others
 ORIGIN

Query Match 4.4%; Score 21; DB 2; Length 122697;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 CATGACATTAAGTATAGCT 151
 |||||||
 Db 100162 CATGACATTAAGTATAGCT 100142

REFERENCE	
AUTHORS	
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckjord,R., Boguslavsky,L., Borkinbatter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., FitzHugh,W., Forrest,C., Funke,R., Gage,D., Horton,L., Galagan,J., Gardyne,S., Grant,G., Haggas,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lehockzy,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meidim,J., Morrow,J.J., Mayor,J., Normand,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollicino,V., Roy,A., Santos,R., Severy,P., Stango-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Turrell,A., Vassiliev,H., Wo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A., and Zody,M.	COMMENT
TITLE	JOURNAL
RESEARCH	Submitted (15-OCT-1999), Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
ALL REPORTS	All reports were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html
CONTACT	Center: Whitehead Institute/MIT Center for Genome Research Center code: WIBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu
PROJECT INFORMATION	Center Project name: L3509 Center clone name: 16_B_11 ----- Summary Statistics Sequencing vector: M13; MT7815; 100% of reads Chemistry: Dye-terminator Big Dye; 100% of reads Assembly program: Phrap; version 0.960731 Consensus quality: 125296 bases at least Q40 Consensus quality: 138924 bases at least Q30 Consensus quality: 146180 bases at least Q20 Insert size: 149388; sum-of-contigs Quality coverage: 8 in Q20 bases; sum-of-contigs Quality coverage: 4.8 in Q20 bases; sum-of-contigs ----- * NOTE: This is a 'working draft' sequence. It currently * consists of 7 contigs. The true order of the pieces * is not known and their order in this sequence record is * arbitrary. Gaps between the contigs are represented as * runs of N, but the exact sizes of the gaps are unknown. * This record will be updated with the finished sequence * as soon as it is available and the accession number will * be preserved. * 1: 2: contig of 2 bp in length * 3: 102: gap of 100 bp * 103: 4217: contig of 4115 bp in length * 4218: 4317: gap of 100 bp * 4318: 11679: contig of 7362 bp in length * 11680: 1179: gap of 100 bp * 11780: 26507: contig of 14728 bp in length * 26508: 26607: gap of 100 bp * 26608: 60280: contig of 33673 bp in length * 60281: 60380: gap of 100 bp * 60381: 103134: contig of 42154 bp in length * 103135: 102334: gap of 100 bp * 103235: 149981: contig of 46754 bp in length. FEATURES source /organism="Homo sapiens" /clone="RP1-16B11" /clone_lid="RPCI-11 Human Male BAC" 1..2 /note="assembly_fragment clone_end:left" vector_side:left"
BASE COUNT	47010 a 27490 c 27742 g 47146 t 600 others
ORIGIN	BASE COUNT Query Match 4.4%; Score 21; DB 9; Length 156248; Best Local Similarity 100.0%; Pred. No. 4; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	241 CTACATAATTAATCAACTTTC 141378 Db 141358 CTACATAATTAATCAACTTTC 261 Oy 241 CTACATAATTAATCAACTTTC 261 /note="assembly_fragment clone_end:SP6 vector_side:right" /note="assembly_fragment clone_end:SP6
RESULT 13	RESULT 13 AC027322/c AC027322 DEFINITION Homo sapiens chromosome 5 clone CTD-2027G10, complete sequence. ACCESSION AC027322 VERSION AC027322.4 GI:17386250 KEYWORDS RTG SOURCE ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. REFERENCE 1 AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center. TITLE Direct Submission JOURNAL Unpublished
REFERENCE	REFERENCE 1 (bases 1 to 15248) DOE Joint Genome Institute and Stanford Human Genome Center. DIRECT SUBMISSION 2 (bases 1 to 15248) REFERENCE 2 (bases 1 to 15248) DOE Joint Genome Institute. DIRECT SUBMISSION 3 (bases 1 to 15248) REFERENCE 3 (bases 1 to 15248) DOE Joint Genome Institute and Stanford Human Genome Center. DIRECT SUBMISSION Submitted (06-DEC-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA On Dec 6, 2001 this sequence version replaced 91:9256705. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www-shgc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.3. FEATURES source Location/Qualifiers 1..156248 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="5" /clone="CTD-2027G10" BASE COUNT 49224 a 29923 c 29730 g 47371 t ORIGIN
QY	241 CTACATAATTAATCAACTTTC 261

Db 124126 CTACATTAATTACAACCTTGTG 124106
 REFSET 14
 AC095224
 LOCUS AC095224 Rattus norvegicus clone CH230-9M19, *** SEQUENCING IN PROGRESS ***,
 DEFINITION 77 unordered pieces.
 ACCESSION AC095224
 VERSION 3 GI:21728983
 KEYWORDS HTG; HTGS PHASEL.
 SOURCE Norway rat.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus;
 REFERENCE 1 (bases 1 to 162227)
 AUTHORS MUZNY,D.M., Adams,C., Adio-Oduola,B., Ali-oustan,F.R., Allen,C.,
 Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
 Barbara,J., Benton,J., Binagle,K., Blanckenburg,K., Bonin,D.,
 Bouck,J., Bowie,S., Blieva,M., Brown,E., Brown,M., Bryant,N.P.,
 Buah,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.G.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chodhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davis,M.L., Davis,C., Davy-Carroll,L., Deederich,D.A.,
 Delaney,K.R., Delgado,O., Dennis,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Duau-Rocha,S., Durbin,K.J.,
 Earhart,C., Edgar,D., Edwards,C.C., Elhaij,C., Escott,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guavara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., Hernández,J.,
 Hernandez,O., Hodges,A., Hogues,M., Holloway,C., Hollins,B.,
 Honsi,F., Howard,S., Huber,J., Hulyk,S., Jackson,L.E.,
 Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
 Karlsson,E., Kelly,S., Khan,U., King,L., Kovari,J., Kovari,C.,
 Kratovic,J., Kuteshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtig,O., Lieu,C., Liu,J., Liu,W., Louis,G.,
 Lledo,R.J., Lu,X., Lucier,A., Lucier,K., Luna,R., Ma,J.,
 Maheshwari,M., Marpaun,A., Martindale,A., Martinez,E.,
 Massey,E., Mawhinney,E., McLeod,M.P., Meddor,M., Mel,G., Metzker,M.,
 Miner,G., Miner,Z., Mitchell,T., Mobabat,K., Morgan,M., Morris,S.,
 Moser,M., Neal,D., Newison,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokekwuo,S., Oguh,M., Okwuonu,G.,
 Oragunye,N., Oviido,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Robas,A., Rojbobulan,T., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shoochani,N., Sison,I.,
 Sodergren,E., Somaik,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabor,P., Tamerska,A., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wlezyk,P., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstock,G. and Gibbs,R.
 TITLE Direct Submission
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 162227)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (16-SEP-2001) Human Genome Sequencing Center; Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 162227)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (12-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 COMMENT On Jul 11, 2002 this sequence version replaced gi:17942456.

Center: Baylor College of Medicine
 Center code: BCM
 Web site: <http://www.hgsc.bcm.tmc.edu/>
 Contact: hgsc-help@bcm.tmc.edu
 ----- Project Information
 Center project name: GDB
 Center clone name: CH230-9M19
 ----- Summary Statistics
 Sequencing vector: Plasmid;
 Chemistry: Dye-terminator Big Dye: 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 95945 bases at least Q40
 Consensus quantity: 100152 bases at least Q30
 Consensus quality: 104634 bases at least Q20
 * NOTE: Estimated insert size may differ from sequence length
 * (see <http://www.hgsc.bcm.tmc.edu/docs/genbank/draft.data.html>).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 77 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 * 1 1013: contig of 1013 bp in length
 * 1014 1113: gap of unknown length
 * 1114 2116: contig of 1003 bp in length
 * 2117 2216: gap of unknown length
 * 2217 3332: contig of 1116 bp in length
 * 3333 3432: gap of unknown length
 * 3432 4854: contig of 1422 bp in length
 * 4855 4954: gap of unknown length
 * 4955 6233: contig of 1279 bp in length
 * 6234 6333: gap of unknown length
 * 6334 7408: contig of 1075 bp in length
 * 7408 7508: gap of unknown length
 * 7509 8753: contig of 1245 bp in length
 * 8754 8853: gap of unknown length
 * 8854 9858: contig of 1005 bp in length
 * 9859 9958: gap of unknown length
 * 9959 11088: contig of 1130 bp in length
 * 11089 11189: gap of unknown length
 * 11189 12543: contig of 1355 bp in length
 * 12544 12643: gap of unknown length
 * 12643 1418: contig of 1475 bp in length
 * 14119 14119: gap of unknown length
 * 14218 14218: gap of unknown length
 * 14219 15223: contig of 1305 bp in length
 * 15223 15623: gap of unknown length
 * 15624 16843: contig of 1219 bp in length
 * 16843 16942: gap of unknown length
 * 16942 18128: contig of 1186 bp in length
 * 18129 18228: gap of unknown length
 * 18228 19293: contig of 1065 bp in length
 * 19293 19393: gap of unknown length
 * 19393 19394: contig of 1094 bp in length
 * 19394 20487: gap of unknown length
 * 20487 20488: gap of unknown length
 * 20488 21906: contig of 1313 bp in length
 * 21906 22006: gap of unknown length
 * 22006 223478: contig of 1472 bp in length
 * 223478 223479: gap of unknown length
 * 223479 23576: gap of unknown length
 * 23576 25150: contig of 1572 bp in length
 * 25150 25250: gap of unknown length
 * 25250 25251: contig of 1841 bp in length
 * 25251 27091: contig of 1841 bp in length
 * 27091 27191: gap of unknown length
 * 27191 27192: contig of 1316 bp in length
 * 27192 29107: contig of 1047 bp in length
 * 29107 29207: gap of unknown length
 * 29207 29208: contig of 1047 bp in length
 * 29208 30255: gap of unknown length
 * 30255 30355: contig of 1122 bp in length
 * 30355 31476: contig of 1122 bp in length
 * 31476 31576: gap of unknown length
 * 31576 32793: contig of 1217 bp in length
 * 32793: gap of unknown length

ACCESSION	AX008577
VERSION	AX008577.1 GI:9996127
KEYWORDS	synthetic construct.
SOURCE	synthetic construct.
ORGANISM	artificial sequences.
REFERENCE	1 (bases 1 to 29)
AUTHORS	Draper,J., Kenton,P. and Paul,W.
TITLE	Inducible Promoters
JOURNAL	Patent: WO 996057-A 14 23-DEC-1999;
DRAPER JOHN (GB); KENTON PAUL (GB); BROGEMMA UK LTD (GB); PAUL	WYATT (GB)
FEATURES	Location/Qualifiers
Source	1..29 /organism="synthetic construct" /db,xref="taxon:32630" /note="primer"
BASE COUNT	10 a 6 c 6 g 7 t
ORIGIN	
Query Match	4.2%; Score 20; DB 6; Length 29;
Best Local Similarity	100.0%; Pred. No. 15; Mismatches 0; Indels 0; Gaps 0;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	226 TTGGAAACTGAACTTACA 245
Db	10 TTGGAAACTGAACTTACA 29
Search completed: July 20, 2003, 06:50:59	
Job time : 1541 secs	
Query Match	4.2%; Score 20; DB 6; Length 29;
Best Local Similarity	100.0%; Pred. No. 15; Mismatches 0; Indels 0; Gaps 0;
Matches	20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	226 TTGGAAACTGAACTTACA 245
Db	10 TTGGAAACTGAACTTACA 29
Search completed: July 20, 2003, 06:50:59	
Job time : 1541 secs	
Query Match	4.2%; Score 20; DB 6; Length 29;
Best Local Similarity	100.0%; Pred. No. 4; Mismatches 0; Indels 0; Gaps 0;
Matches	21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	250 TTCAACTTTCGAAATAAA 270
Db	63077 TTCAACTTTCGAAATAAA 63097

RESULT 15

AX008577

LOCUS

DEFINITION

Sequence 14 from Patent WO996057.

DNA

linear

PAT 06-SEP-2000

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2003, 05:20:09 ; Search time 189 Seconds
 (without alignments)

16559.784 Million cell updates/sec

Title: US-09-719-002-1

Perfect score: 1

Sequence: gaatctttatgcgactga.....acaacacaaacaaatcatg 475

Scoring table: OLIGO_NTC

Word size : 0

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_101002,*

1: /SIDS2/geodata/geneseq/geneseq/geneseq-emb1/NA1980.DAT:*

2: /SIDS2/geodata/geneseq/geneseq-emb1/NA1981.DAT:*

3: /SIDS2/geodata/geneseq/geneseq-emb1/NA1982.DAT:*

4: /SIDS2/geodata/geneseq/geneseq-emb1/NA1983.DAT:*

5: /SIDS2/geodata/geneseq/geneseq-emb1/NA1984.DAT:*

6: /SIDS2/geodata/geneseq/geneseq-emb1/NA1985.DAT:*

7: /SIDS2/geodata/geneseq/geneseq-emb1/NA1986.DAT:*

8: /SIDS2/geodata/geneseq/geneseq-emb1/NA1987.DAT:*

9: /SIDS2/geodata/geneseq/geneseq-emb1/NA1988.DAT:*

10: /SIDS2/geodata/geneseq/geneseq-emb1/NA1989.DAT:*

11: /SIDS2/geodata/geneseq/geneseq-emb1/NA1990.DAT:*

12: /SIDS2/geodata/geneseq/geneseq-emb1/NA1991.DAT:*

13: /SIDS2/geodata/geneseq/geneseq-emb1/NA1992.DAT:*

14: /SIDS2/geodata/geneseq/geneseq-emb1/NA1993.DAT:*

15: /SIDS2/geodata/geneseq/geneseq-emb1/NA1994.DAT:*

16: /SIDS2/geodata/geneseq/geneseq-emb1/NA1995.DAT:*

17: /SIDS2/geodata/geneseq/geneseq-emb1/NA1996.DAT:*

18: /SIDS2/geodata/geneseq/geneseq-emb1/NA1997.DAT:*

19: /SIDS2/geodata/geneseq/geneseq-emb1/NA1998.DAT:*

20: /SIDS2/geodata/geneseq/geneseq-emb1/NA1999.DAT:*

21: /SIDS2/geodata/geneseq/geneseq-emb1/NA2000.DAT:*

22: /SIDS2/geodata/geneseq/geneseq-emb1/NA2001A.DAT:*

23: /SIDS2/geodata/geneseq/geneseq-emb1/NA2001B.DAT:*

24: /SIDS2/geodata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	475	100.0	475	21	A2Z29510
2	24	5.1	42	21	A2Z29523
3	22	4.5	29	21	A2Z29516
4	4.6	3.3	21	A2Z29514	
5	22	4.6	40	21	A2Z29519
6	22	4.6	40	21	A2Z29521
7	20	4.2	29	21	A2Z29517
8	20	4.2	29	21	A2Z29518
9	20	4.2	41	21	A2Z29515

ALIGNMENTS

10	19	4.0	44	21	A2Z29520	Primer-1 used for
11	19	4.0	44	21	A2Z29522	Primer-3 used for
12	19	4.0	47	22	A2H28638	Drosophila melanog
13	19	4.0	507	21	AAC43329	Arabidopsis thalia
14	19	4.0	1143	22	AAF63492	Drosophila gustato
15	19	4.0	1161	21	AAC3610	Arabidopsis thalia
16	19	4.0	1535	21	AAC41107	Arabidopsis thalia
17	19	4.0	14568	24	ABL2231	Human immune syste
18	19	4.0	28911	23	ABL20204	Drosophila melanog
19	19	4.0	25929	24	ABV25595	Drosophila melanog
20	18	3.8	397	24	ABK3503	Rat sequence diffe
21	18	3.8	487	22	AAB18300	Human breast speci
22	18	3.8	529	24	ABQ24226	Oligonucleotide fo
23	18	3.8	529	24	ABQ24227	Oligonucleotide fo
24	18	3.8	860	23	AAB9834	Human EST-derived
25	18	3.8	1205	23	ABY2559	Human prostate exp
26	18	3.8	1492	24	ABK93442	Human CDNA encodin
27	18	3.8	1609	24	ABK93446	Drosophila melanog
28	18	3.8	2730	23	ABL2246	Drosophila melanog
29	18	3.8	3215	23	ABL1002	Drosophila melanog
30	18	3.8	3420	23	ABL10200	Drosophila melanog
31	18	3.8	5998	24	AB133135	Human immune syste
32	18	3.8	6973	24	ABL33778	Human gene signatu
33	17	3.6	68	15	ABT1824	Porcine circovirus
34	17	3.6	240	20	AAX83156	Human ORF polynuc
35	17	3.6	253	24	ABN21432	Human colon cancer
36	17	3.6	300	21	AA00270	Human ORF polynuc
37	17	3.6	378	24	ABN21444	Human G protein-co
38	17	3.6	392	22	ABX46855	CDNA encoding nove
39	17	3.6	392	24	ABX81533	Human prostate exp
40	17	3.6	398	23	ABV03703	S. epidermidis ope
41	17	3.6	429	22	AAB53822	Staphylococcus epi
42	17	3.6	447	24	ABN9120	Human secreted pro
43	17	3.6	448	21	ACG5146	CDNA encoding lung
44	17	3.6	482	24	ABR38805	Human G protein-co
45	17	3.6	494	22	ABX46858	

XX Asparagus officinalis.

XX Key location/Qualifiers

XX misc_feature

XX misc_feature</

FT /*tag= d
 FT /note= "Homologous to Carrot PR-3 and PR-4 promoter
 FT sequence"
 FT repeat_region
 FT 281..318
 FT /*tag= e
 FT /rp-type= TANDEM
 FT 281..298
 FT /*tag= f
 FT /note= "18 bp repeat"
 FT 409..416
 FT /*tag= g
 FT XX
 PN WO966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PP 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 PI Draper, J., Kenton, P., Paul, W;
 XX
 DR WPI; 2000-106107/09.
 XX
 PT Novel Promoters used to control the expression of heterologous genes in
 transformed plants -
 PS
 Claim 4; Fig 6; 67pp; English.
 XX
 CC The present DNA sequence is a novel inducible promoter, derived from
 CC Asparagus officinalis thiamatin-like PR-5 related gene, AOPRT-L. The
 CC promoter sequence has similarities with other PR promoters. The
 CC and non-phytotoxic inducing agents, like Salicylic acid or BTH. The
 CC promoters also exhibit low levels of pathogen induced systemic
 CC activation and environmentally induced expression.
 CC The inducible promoter sequence is used to control the expression of
 CC heterologous genes in transformed plants, especially genes whose
 CC products affect a trait of the plant, such as pathogen resistance,
 CC disease control, sterility, fertility or fruit ripening.
 XX
 Sequence 475 BP; 163 A; 113 C; 67 G; 132 T; 0 other;
 SQ
 Query Match 100 %; Score 475; DB 21; Length 475;
 Best Local Similarity 100.0%; Pred. No. 7.7e-236;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY
 1 GATTTCTTATGGCCTGACTCTCTGTGCTGCCGAGGTCGTCGAAATTTCGT 60
 DB 1 GAATTCTATGGCCTGACTCTCTGTGCTGCCGAGGTCGTCGAAATTTCGT 60
 QY
 61 TGGCGACACATACGGCCCTGCTTGATTTGACAGTCATATAATTCCATGTC 120
 DB 61 TGGCGACACATACGGCCCTGCTTGATTTGACAGTCATATAATTCCATGTC 120
 QY
 121 GAGAGAGACATGACTTAATGTAATAGCTTCAATAAATAGAAGTCTAACGTGAGCT 180
 DB 121 GAGAGAGACATGACTTAATGTAATAGCTTCAATAAATAGAAGTCTAACGTGAGCT 180
 QY
 181 GACACTTCAGAGAAAATCTAACTTAACTGCTTGGCTTGAATTGAAACTGATC 240
 DB 181 GACACTTCAGAGAAAATCTAACTTAACTGCTTGAATTGAAACTGATC 240
 QY
 241 CTACATTAATCACACTTGCATAATAAATAGAAGTCTAACGTGAGCT 300
 DB 241 CTACATTAATCACACTTGCATAATAAATAGAAGTCTAACGTGAGCT 300
 QY
 301 TCTAACATGAGACTGTCAGAAGTCTGAGACTTCA 360
 DB 301 TCTAACATGAGACTGTCAGAAGTCTGAGACTTCA 360

QY 361 CAATGGAGATATCCATGGACTGATGACACCAATTATCCATAATACCTG 420
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 361 CAATGGAGATATCCATGGACTGATGACACCAATTATCCATAATACCTG 420
 QY 421 CCCATCCCTCTCAGACTCATCTACTCAAACACACACAACTCATG 475
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 Db 421 CCCATCCCTCTCAGACTCATCTACTCAAACACACACAACTCATG 475
 RESULT 2
 AAZ2953/C
 ID AAZ2953 standard; DNA; 42 BP.
 XX
 AC AAZ29523;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-4 used for construction of AOPRT-Lx3 promoter.
 XX
 KW Inducible Promoter; Thiamatin-like PR-5 related gene; AOPRT-L; Primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; AOPRT-Lx3 promoter; p22-JIT60; pUC19; ss.
 XX
 OS Synthetic.
 XX
 PN WO966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PR 21-JUN-1999; 99WO-GB01949.
 XX
 PA 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 DR Draper, J., Kenton, P., Paul, W;
 XX
 WPI; 2000-106107/09.
 XX
 PT Novel Promoters used to control the expression of heterologous genes in
 transformed plants -
 PS
 Example 12; Page 41; 67pp; English.
 XX
 CC The present DNA sequence is a PCR primer-4, used for the construction of
 CC the AOPRT-Lx3 promoter. This primer is used to amplify the region from
 CC -133 bp to -247 bp of the AOPRT-L promoter, from p22-JIT60 and cloned
 CC into pUC19.
 XX
 Sequence 42 BP; 13 A; 7 C; 14 G; 8 T; 0 other;
 SQ
 Query Match 5.1%; Score 24; DB 21; Length 42;
 Best Local Similarity 100.0%; Pred. No. 0.028; Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY
 317 GTCCACGAACTGTACCTTATTC 340
 DB 42 GTCACGAACTGTACCTTATTC 19

RESULT 3
 AAZ29516
 AAZ29516 standard; DNA; 29 BP.
 XX
 AC AAZ29516;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-1 for identification of SA responsive element in AOPRT-L promoter.
 XX
 KW Inducible Promoter; Thiamatin-like PR-5 related gene; AOPRT-L; Primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; environmental;

KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; Inverse PCR; IPCR; ss.
 XX
 OS Synthetic.
 PN WO9966057-A2.
 XX 23-DEC-1999.
 PD
 PF 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX (BIOG-) BIOGEMMA UK LTD.
 PA Draper J, Kenton P, Paul W;
 PI DR
 XX DR
 XX WPI; 2000-106107-09.
 PT Novel Promoters used to control the expression of heterologous genes in
 PR transformed plants
 XX Example 12: Page 40; 67pp; English.
 PS
 P2 The present DNA sequence is a PCR primer-1, used for the identification
 CC and multimerisation of a salicylic acid, SA/BTH responsive element in
 CC the AOPRT-L promoter region. This primer is designed against both the 5' and 3' ends
 CC of the promoter, with extensions to provide appropriate restriction
 CC sites for further cloning.
 XX Sequence 33 BP; 6 A; 10 C; 7 G; 10 T; 0 other;
 SQ Query Match 4.6%; Score 22; DB 21; Length 33;
 KW Best Local Similarity 100.0%; Pred. No. 0; 3; Mismatches 0; Indels 0; gaps 0;
 CC Matches 22; Conservative 0; Mismatches 0; Indels 0; gaps 0;
 CC of GUS fusion constructs.
 XX Sequence 29 BP; 9 A; 7 C; 8 G; 5 T; 0 other;
 SQ Query Match 4.6%; Score 22; DB 21; Length 29;
 KW Best Local Similarity 100.0%; Pred. No. 0; 3; Mismatches 0; Indels 0; Gaps 0;
 CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 110 TCCGATCTCAGAGAGAGCAC 131
 DB 8 TTCCATGTCATGAGAGAGCAC 29
 RESULT 4
 AAZ29514 AAZ29514 standard; DNA; 33 BP.
 XX
 AC AAZ29514;
 XX
 DT 14-MAR-2000 (first entry)
 DE Primer-4 for identification of SA responsive element in AOPRT-L promoter.
 XX Inducible promoter; Traumatin-like PR-5 related gene; AOPRT-L; primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; template; p22-JIT60; Inverse PCR; IPCR; ss.
 OS Synthetic.
 XX WO9966057-A2.
 PN
 XX
 PD 23-DEC-1999.
 PR
 XX 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX (BIOG-) BIOGEMMA UK LTD.
 PA Draper J, Kenton P, Paul W;
 PI DR
 XX DR
 XX WPI; 2000-106107-09.
 PT Novel Promoters used to control the expression of heterologous genes in
 PR transformed plants
 XX Example 12: Page 40; 67pp; English.
 PS The present DNA sequence is a PCR primer-4, used for the identification
 CC and multimerisation of a salicylic acid, SA/BTH responsive element in
 CC the AOPRT-L promoter region. This primer is used together with PCR
 CC primers 1-3, using p22-JIT60 as template, for the construction of
 CC GUS fusion constructs.
 XX Sequence 40 BP; 4 A; 6 C; 14 G; 16 T; 0 other;
 SQ Query Match 4.6%; Score 22; DB 21; Length 40;
 KW Best Local Similarity 100.0%; Pred. No. 0; 3; Mismatches 0; Indels 0; Gaps 0;
 CC Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 451 CAACACACACACACCACTC 472
 DB 40 CAAACACACACACACCACTC 19
 PT Novel Promoters used to control the expression of heterologous genes in
 PR transformed plants - to control the expression of heterologous genes in

RESULT 6
 AA229521/C
 ID AA229521 standard; DNA: 40 BP.
 XX
 AC AA229521;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE - Primer-2 used for construction of AOPRT-Lx3 promoter.
 XX
 KW Inducible promoter; Thaumatin-like PR-5 related gene; AOPRT-L; primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; AOPRT-Lx3 promoter; p22-JIT60; pUC19; ss.
 XX
 OS Synthetic.
 XX
 PN WO9966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-GB01949.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 PI Draper J, Kenton P, Paul W;
 XX
 DR WPI; 2000-106107/09.
 XX
 PT Novel promoters used to control the expression of heterologous genes in
 PT transformed plants -
 XX
 PS Example 12; Page 40; 67pp; English.
 XX
 CC The present DNA sequence is a PCR Primer-2, used for the identification
 CC and multimerisation of a salicylic acid, SA/BTH responsive element in
 CC the AOPRT-L promoter region. This primer is designed to regions of
 CC AOPRT-L promoter and used along with PCR primer-4 for the construction
 CC of GUS fusion constructs.
 XX
 SQ Sequence 29 BP; 10 A; 6 C; 6 G; 7 T; 0 other;
 PR Best Local Similarity 4.2%; Score 20; DB 21; Length 29;
 PR Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 PR Qy 226 TTTGAACTGTAACTCTACA 245
 DR Db 10 TTTGAACTGTAACTCTACA 29

RESULT 7
 AA229517
 ID AA229517 standard; DNA: 29 BP.
 XX
 AC AA229517;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-3 for identification of SA responsive element in AOPRT-L promoter.
 XX
 KW Inducible promoter; Thaumatin-like PR-5 related gene; AOPRT-L; primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; Inverse PCR; IPCR; ss.
 XX
 OS Synthetic.
 XX
 PN WO9966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PF 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 PI Draper J, Kenton P, Paul W;
 XX
 DR WPI; 2000-106107/09.
 XX
 PT Novel promoters used to control the expression of heterologous genes in
 PT transformed plants -
 XX
 PS Example 12; Page 40; 67pp; English.
 XX
 OS The present DNA sequence is a PCR primer-3, used for the identification
 CC and multimerisation of a salicylic acid, SA/BTH responsive element in
 CC the AOPRT-L promoter region. This primer is designed to regions of
 CC AOPRT-L promoter and used along with PCR Primer-4 for the construction
 CC of GUS fusion constructs.

SQ Sequence 29 BP; 9 A; 7 C; 6 G; 7 T; 0 other;
 Query Match 4.2%; Score 20; DB 21; Length 29;
 Best Local Similarity 100.0%; Pred. No. 3.3%; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0
 RESULT 9
 AAZZ9515/^C
 ID AAZZ9515 standard; DNA; 41 BP.
 XX
 AC AAZZ9515;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-2 for construction of AOPRT-L promoter-GUS chimeric gene.
 XX
 KW Inducible promoter; Thaumatin-like PR-5 related gene; AOPRT-L; primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTI; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; AOPRT-Lx3 promoter; p22-JTn60; PUC19; ss.
 XX
 OS Synthetic.
 XX
 PN WO9966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PI Draper J, Kenton P, Paul W;
 XX
 DR 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 PS Example 12; Page 41; 67pp; English.
 XX
 CC The present DNA sequence is a PCR Primer-1' used for the construction of
 CC the AOPRT-Lx3 promoter. This primer is used to amplify the region from
 CC 0 bp to -247 bp of the AOPRT-L promoter, from p22-JTn60 and cloned into
 CC pUC19.
 XX
 SQ Sequence 44 BP; 10 A; 11 C; 10 G; 13 T; 0 other;
 Query Match 4.0%; Score 19; DB 21; Length 44;
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 11
 AAZZ9522
 ID AAZZ9522 standard; DNA; 44 BP.
 XX
 AC AAZZ9522;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-3 used for construction of AOPRT-Lx3 promoter.
 XX
 KW Inducible promoter; Thaumatin-like PR-5 related gene; AOPRT-L; primer;
 KW non-phytotoxic inducing agent; Salicylic acid; SA; BTI; environmental;
 KW developmental; GUS construct; multimerisation; SA responsive element;
 KW systemic activation; AOPRT-Lx3 promoter; p22-JTn60; pUC19; ss.
 XX
 OS Synthetic.
 XX
 PN WO9966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 PR 21-JUN-1999; 99WO-GB01949.
 XX
 PR 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA JK LTD.
 XX
 PI Draper J, Kenton P, Paul W;
 XX
 DR WPI; 2000-106107/09.
 XX
 PS Example 3; Page 32; 67pp; English.
 XX
 CC The present DNA sequence is a PCR Primer-2, used for the isolation of
 CC the AOPRT-L promoter region. The promoter sequence was obtained from pICR-TA
 CC using this primer. This primer is designed against both the 5' and 3' ends
 CC of the promoter, with extensions to provide appropriate restriction
 CC sites for further cloning.
 XX
 SQ Sequence 41 BP; 4 A; 8 C; 14 G; 15 T; 0 other;
 Query Match 4.2%; Score 20; DB 21; Length 41;
 Best Local Similarity 100.0%; Pred. No. 3.3%; Mismatches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 RESULT 10
 AAZZ9520
 ID AAZZ9520 standard; DNA; 44 BP.
 XX
 AC AAZZ9520;
 XX
 DT 14-MAR-2000 (first entry)
 XX
 DE Primer-1 used for construction of AOPRT-Lx3 promoter.
 XX

XX Novel promoters used to control the expression of heterologous genes in
 PT transformed plants
 XX

* Example 12; Page 41; 67pp; English.

XX The present DNA sequence is a PCR primer-3' used for the construction of
 CC the AcPRT-Lx3 promoter. This primer is used to amplify the region from
 CC -113 bp to -247 bp of the AcPRT-L promoter, from p22-JIT60 and cloned
 CC into pGK19.
 XX

Sequence 44 BP; 10 A; 11 C; 10 G; 13 T; 0 other;
 Query Match 4.0%; Score 19; DB 21; Length 44;
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 19; Conservative 0; Indels 0; Gaps 0;
 Matches 19; Query 226 TTGGAACTGAAACCTAG 244
 Db 26 TTGGAACTGAAACCTAG 44

RESULT 12
 ID AAH2963B/C
 XX AAH2963B standard; DNA; 47 BP.
 AC XX
 XX DT 17-JUL-2001 (first entry)
 XX Drosophila melanogaster essential gene fragment, SEQ ID NO: 827.
 XX KW Drosophila melanogaster; fruit fly; essential gene; screening assay;
 KW pesticide; crop protection; chromosome 2; ds.
 OS Drosophila melanogaster.
 XX WO200118547-A1.
 XX PR 07-SEP-1999; 99GB-0021009.
 PD 15-MAR-2001.
 PP 06-SEP-2000; 2000WO-GB03444.
 XX PR 07-SEP-1999; 99GB-0021009.
 PA (UNIU) UNIV GLASGOW.
 PI Davies RW, Kaiser K, Yang MY;
 XX DR WPI; 2001-281436/29.
 XX PT Screening assays for identifying compounds having a physiological effect on proteins identified as being essential -
 PT XX Claim 1; Page 642; 695pp; English.
 PS XX
 XX The present sequence is part of an essential gene from Drosophila melanogaster. Lack of expression of the protein encoded by this gene leads to a lethal or semi-lethal phenotype. The invention relates to 902 nucleic acid sequences from genes encoding proteins which are thought to be essential, and to a screening assay for identifying compounds which have a physiological effect on these proteins. Suitable compounds are useful as pesticides and may be used in conjunction with other pesticides and herbicides for crop protection. The gene corresponding to the present sequence is located on chromosome 2.
 Sequence 47 BP; 13 A; 6 C; 9 G; 19 T; 0 other;
 Query Match 4.0%; Score 19; DB 22; Length 47;
 Best Local Similarity 100.0%; Pred. No. 11; Mismatches 19; Conservative 0; Indels 0; Gaps 0;
 Matches 19; Query 249 ATTACACTTGTGAAATA 267
 Db 41 ATTACACTTGTGAAATA 23

RESULT 13
 ID AAC44332/C
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 42470.
 XX KW Hybridisation assay; genetic mapping; gene expression control; protein identification; signal transduction pathway; metabolic pathway; promoter; termination sequence; ss.
 KW OS Arabidopsis thaliana.
 PN XX EP1033405-A2.
 XX PD 06-SEP-2000.
 XX PF 25-FEB-2000; 2000EP-0301439.
 XX PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123584.
 PR 23-MAR-1999; 99US-0126264.
 PR 25-MAR-1999; 99US-0126785.
 PR 29-MAR-1999; 99US-0127462.
 PR 01-APR-1999; 99US-0128234.
 PR 06-APR-1999; 99US-0128714.
 PR 08-APR-1999; 99US-0129714.
 PR 16-APR-1999; 99US-0130077.
 PR 19-APR-1999; 99US-0130449.
 PR 21-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 30-APR-1999; 99US-0134007.
 PR 04-MAY-1999; 99US-0134484.
 PR 05-MAY-1999; 99US-013485.
 PR 06-MAY-1999; 99US-013486.
 PR 06-MAY-1999; 99US-013487.
 PR 07-MAY-1999; 99US-013488.
 PR 07-MAY-1999; 99US-013489.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134270.
 PR 18-MAY-1999; 99US-0134668.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-013353.
 PR 24-MAY-1999; 99US-0136229.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136592.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137702.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-013840.
 PR 14-JUN-1999; 99US-013847.
 PR 16-JUN-1999; 99US-013919.
 PR 16-JUN-1999; 99US-013952.
 PR 17-JUN-1999; 99US-013953.

ID AAF53749 standard; DNA; 1143 BP.
 XX
 AC AAF63749;
 XX
 -DT 03-APR-2001 (first entry)
 DE Drosophila gustatory receptor GR39D.2c DNA sequence.
 XX - Gustatory receptor; fruit fly; taste; pheromone; semiochemical;
 KW - crop damage; pest control; ds;
 XX OS Drosophila melanogaster.
 XX
 WO20007208-A2.
 PN
 XX
 PD 21-DEC-2000.
 XX
 PF 14-JUN-2000; 2000WO-US16211..
 XX
 PR 14-JUN-1999; 99US-0138668.
 PR 10-FEB-2000; 2000US-0181704.
 XX
 PA (UYYA) UNIV YALE.
 XX
 PT Carlson PJ, Clyne PJ, Warr CG;
 XX
 DR WPI; 2001-061873/07.
 DR P-FSDB; AAB75210.
 XX
 PT New isolated nucleic acid molecule encoding Drosophila Gustatory Receptor protein useful for e.g. identification of compounds which may be used for pest management.
 XX
 PS Claim 3; Page 134-135; 227pp; English.
 XX
 CC This invention relates to polynucleotide sequences AAF53732 & AAF63777 which encode Drosophila gustatory receptor proteins represented by sequences AAB75193 - AAB75238. The invention includes methods for determining gustatory receptor ligands. Also included is a method for modulating the expression of the DNA encoding the receptors. The DNA and protein sequences may be used for the identification of compounds, e.g. pheromones and other semiochemicals, which may be used for pest management. The DNA sequences may also be used for behavioural studies involving gustatory systems in various organisms. Also, the DNA sequences may also be used to track down gustatory receptor genes in insects that damage crops or transmit diseases.
 CC
 SQ sequence 1143 BP; 326 A; 232 C; 226 G; 359 T; 0 other;
 Query Match 4.0%; Score 19; DB 22; Length 1143;
 Best Local Similarity 100%; Pred. No. 10;
 Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 249 ATTACACTTGTCAATA 267
 Db 641 ATTACACTTGTCAATA 659
 RESULT 15
 AAC36210/C
 ID AAC36210 standard; DNA; 1161 BP.
 XX
 AC AAC36210;
 XX
 DT 17-OCT-2000 (first entry)
 DE Arabidopsis thaliana DNA fragment SEQ ID NO: 12948.
 KW Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.
 XX
 OS Arabidopsis thaliana.

XX EP1033405-A2.
 PN
 XX
 PD 06-SEP-2000.
 XX
 PP 25-FEB-2000; 2000EP-0301439.
 XX
 PR 25-FEB-1999; 99US-0121825.
 PR 05-MAR-1999; 99US-0123180.
 PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 23-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 05-APR-1999; 99US-0127234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0128945.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 23-APR-1999; 99US-0130891.
 PR 28-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132863.
 PR 11-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139492.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139450.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139467.
 PR 18-JUN-1999; 99US-0139468.
 PR 18-JUN-1999; 99US-0139469.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140553.
 PR 24-JUN-1999; 99US-0140554.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.

PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	05-JUL-1999;	99US-0142200.	PR	23-SEP-1999;	99US-0155886.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155959.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156558.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-015696.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157665.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158332.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158669.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159093.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159395.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159395.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159229.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144884.	PR	14-OCT-1999;	99US-0159337.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159337.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-015984.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145089.	PR	25-OCT-1999;	99US-0160704.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160880.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160767.
PR	26-JUL-1999;	99US-0145276.	PR	25-OCT-1999;	99US-0160704.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161005.
PR	27-JUL-1999;	99US-0145918.	PR	26-OCT-1999;	99US-0161559.
PR	27-JUL-1999;	99US-0145919.	PR	26-OCT-1999;	99US-0161660.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146386.	PR	28-OCT-1999;	99US-0161220.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161192.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161193.
PR	04-AUG-1999;	99US-0147302.	PR	29-OCT-1999;	99US-0162142.
PR	05-AUG-1999;	99US-0147192.			
PR	05-AUG-1999;	99US-0147260.			
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.			
PR	09-AUG-1999;	99US-0147935.			
PR	10-AUG-1999;	99US-0148171.			
PR	11-AUG-1999;	99US-0148319.			
PR	12-AUG-1999;	99US-0148341.			
PR	13-AUG-1999;	99US-0148565.			
PR	16-AUG-1999;	99US-0149368.			
PR	17-AUG-1999;	99US-0149375.			
PR	18-AUG-1999;	99US-0149426.			
PR	20-AUG-1999;	99US-0149722.			
PR	20-AUG-1999;	99US-0149723.			
PR	20-AUG-1999;	99US-0149929.			
PR	23-AUG-1999;	99US-0149902.			
PR	*				
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
*PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	13-SEP-1999;	99US-0153758.			
PR	15-SEP-1999;	99US-0154018.			

Query Match 4.0%; Score 19; DB 21; Length 1161;
 Best Local Similarity 100.0%; Pred No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 19; Conservative 0; Mismatches 0;

Qy 11 TCGCAACTGACTCTCTCTGT 29
 ||||||| ||||| ||||| |||||
 Db 370 TCGCACTGACTCTCTGT 352

Search completed: July 20, 2003, 06:25:08
 Job time : 190 secs

RESULT 3
US-09-347-594-2/C
; Sequence 2, Application US/09347594
; Patent No. 6217883
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon M.
; APPLICANT: MEERAN, Brian M.
; APPLICANT: ELLIS, John A.
; APPLICANT: KRAKOWIAK, George S.
; APPLICANT: AUDONNET, Jean-Christophe F.
; TITLE OF INVENTION: PORCINE CIRCOVIRUS AND PARVOVIRUS VACCINE
; FILE REFERENCE: 454313-2338
; CURRENT APPLICATION NUMBER: US/09/347,594
; CURRENT FILING DATE: 1999-07-01
; EARLIER APPLICATION NUMBER: 98 08777
; EARLIER FILING DATE: 1998-07-06
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Porcine circovirus
US-09-347-594-2

Query Match 3.6%; Score 17; DB 4; Length 1767;
Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY ||||||| 31 GTGCTGCCGAGGTGCTG 47
Db 385 GTGCTGCCGAGGTGCTG 369

RESULT 4
US-09-082-558-1/C
; Sequence 1, Application US/09082558A
; Patent No. 6368601
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEERAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HAASRD, Lori
; APPLICANT: HAROLD, John
; APPLICANT: CHARREIRE, Catherine E.
; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/082,558A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: FR 9800873
; EARLIER FILING DATE: 1998-01-22
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Porcine circovirus
US-09-082-558-2

Query Match 3.6%; Score 17; DB 4; Length 1767;
Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY ||||||| 31 GTGCTGCCGAGGTGCTG 47
Db 385 GTGCTGCCGAGGTGCTG 369

RESULT 5
US-09-082-558-2/C
; Sequence 2, Application US/09082558A
; Patent No. 6368601
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEERAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HAASRD, Lori
; APPLICANT: HAROLD, John
; APPLICANT: CHARREIRE, Catherine E.
; TITLE OF INVENTION: REAGENTS
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/082,558A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: FR 9800873
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: FR 9803707
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: FR 97/12382
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Porcine circovirus
US-09-082-558-2

Query Match 3.6%; Score 17; DB 4; Length 1767;
Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY ||||||| 31 GTGCTGCCGAGGTGCTG 47
Db 385 GTGCTGCCGAGGTGCTG 369

RESULT 6
US-09-161-092-1/C
; Sequence 1, Application US/09161092
; Patent No. 6391314
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEERAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HAASRD, Lori
; APPLICANT: HARDING, John
; APPLICANT: CHARREIRE, Catherine E.
; APPLICANT: CHAPPUIS, Gilles E.
; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/082,558A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: FR 9800873
; EARLIER FILING DATE: 1998-01-22
; EARLIER APPLICATION NUMBER: FR 9803707
; EARLIER FILING DATE: 1998-03-20
; EARLIER APPLICATION NUMBER: FR 97/12382
; EARLIER FILING DATE: 1997-10-03
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; CURRENT APPLICATION NUMBER: US/09/151,092

CURRENT FILING DATE: 1998-09-25
; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/082,558
; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-05-21
; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9800873
; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-01-22
; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 98033707
; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-03-20
; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 97/12382
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 1
; LENGTH: 1767
; TYPE: DNA
; ORGANISM: Porcine circovirus
; US-09-161-092-1

Query Match 3.6%; Score 17; DB 4; Length 1767;
; Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 GTGCTGCCGAGGTGCG 47
Db 385 GTGCCTGCCGGGGTGC 369

RESULT 7
US-09-161-092-2/C

; Sequence 2, Application US/09161092
; General Information:
; Patent No. 6391314
; Applicant: ALLAN, Gordon
; Applicant: MEEHAN, Brian
; Applicant: CLARK, Edward
; Applicant: HAINES, Deborah
; Applicant: HASSARD, Lori
; Applicant: HARDING, John
; Applicant: CHARREIRE, Catherine E.
; Applicant: CHAPPUIS, Gilles E.
; Title of Invention: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; File Reference: ALLAN
; Current Application Number: US/09/161,092
; Current Filing Date: 1998-09-25
; Prior Application Number: EARLIER APPLICATION NUMBER: 09/082,558
; Prior Filing Date: EARLIER FILING DATE: 1998-05-21
; Prior Application Number: EARLIER APPLICATION NUMBER: FR 9800873
; Prior Filing Date: EARLIER FILING DATE: 1998-01-22
; Prior Application Number: EARLIER APPLICATION NUMBER: FR 98033707
; Prior Filing Date: EARLIER FILING DATE: 1998-03-20
; Prior Application Number: EARLIER APPLICATION NUMBER: FR 97/12382
; Prior Filing Date: EARLIER FILING DATE: 1997-10-03
; Number of Seq ID Nos: 6
; Software: Patentin Ver. 2.0
; Seq ID No: 2
; Length: 1767
; Type: DNA
; Organism: Porcine circovirus
; US-09-161-092-2

Query Match 3.6%; Score 17; DB 4; Length 1767;
; Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 GTGCTGCCGAGGTGCG 47
Db 385 GTGCCTGCCGGGGTGC 369

RESULT 9
US-09-347-594-4/C

; Sequence 4, Application US/09347594
; Patent No. 6217883
; General Information:
; Applicant: ALLAN, Gordon M.
; Applicant: MEEHAN, Brian M.
; Applicant: ELIS, John A.
; Applicant: KRAKOWKA, George S.
; Applicant: AUDONET, Jean-Christophe F.
; Title of Invention: PORCINE CIRCOVIRUS AND PARVOVIRUS VACCINE
; File Reference: 454313-2338
; Current Application Number: US/09/347,594
; Current Filing Date: 1999-07-01
; Earlier Application Number: 98 08777
; Earlier Filing Date: 1998-07-06
; Number of Seq ID Nos: 5
; Software: Patentin Ver. 2.1
; Seq ID No: 4
; Length: 1768
; Type: DNA
; Organism: Porcine circovirus
; US-09-347-594-4

Query Match 3.6%; Score 17; DB 4; Length 1768;
; Best Local Similarity 100.0%; Pred. No. 29; Mismatches 0; Indels 0; Gaps 0;
; Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 31 GTGCTGCCGAGGTGCG 47
Db 385 GTGCCTGCCGGGGTGC 369

RESULT 10
US-09-082-558-3/C

; Sequence 3, Application US/09082558A
; Patent No. 6365601
; General Information:
; Applicant: ALLAN, Gordon
; Applicant: MEEHAN, Brian
; Applicant: CLARK, Edward
; Applicant: HAINES, Deborah
; Applicant: HASSARD, Lori
; Applicant: HARDING, John

RESULT 8
US-09-347-594-3/C
; Sequence 3, Application US/09347594
; Patent No. 6217883

APPLICANT: CHARREYRE, Catherine E.
 APPLICANT: CHAPPUIS, Gilles E.
 TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
 FILE REFERENCE: ALLAN
 CURRENT APPLICATION NUMBER: US/09/082,558A
 CURRENT FILING DATE: 1998-05-21
 EARLIER APPLICATION NUMBER: FR 9800873
 EARLIER FILING DATE: 1998-01-22
 EARLIER APPLICATION NUMBER: FR 9803707
 EARLIER FILING DATE: 1998-03-20
 EARLIER APPLICATION NUMBER: FR 97/12382
 EARLIER FILING DATE: 1997-10-03
 NUMBER OF SEQ ID NOS: 6
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 3
 LENGTH: 1768
 TYPE: DNA
 ORGANISM: Porcine circovirus
 US-09-082-558-3.

RESULT 11
 Query Match 3.6%; Score 17; DB 4; Length: 1768;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO 3
 Qy 31 GTGCTGCCGAGGTCTG 47
 |||||||
 Db 385 GTGCTGCCGAGGTCTG 369

US-09-082-558-4/c
 Sequence 4, Application US/09082558A
 ; Parent No. 6368601
 ; GENERAL INFORMATION:
 ; APPLICANT: ALLAN, Gordon
 ; APPLICANT: MEEHAN, Brian
 ; APPLICANT: CLARK, Edward
 ; APPLICANT: HAINES, Deborah
 ; APPLICANT: HASSARD, Lori
 ; APPLICANT: HARDING, John
 ; APPLICANT: CHARREYRE, Catherine E.
 ; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
 ; TITLE OF INVENTION: REAGENTS
 ; FILE REFERENCE: ALLAN
 ; CURRENT APPLICATION NUMBER: US/09/082,558A
 ; CURRENT FILING DATE: 1998-05-21
 ; EARLIER APPLICATION NUMBER: FR 9800873
 ; EARLIER FILING DATE: 1998-01-22
 ; EARLIER APPLICATION NUMBER: FR 9803707
 ; EARLIER FILING DATE: 1998-03-20
 ; EARLIER APPLICATION NUMBER: FR 97/12382
 ; EARLIER FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 6
 ; SEQ ID NO 4
 ; LENGTH: 1768
 ; TYPE: DNA
 ; ORGANISM: Porcine circovirus
 ; OTHER INFORMATION: N represents A or C or G or T
 ; US-09-082-558-4

RESULT 13
 Query Match 3.6%; Score 17; DB 4; Length 1768;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO 3
 Qy 31 GTGCTGCCGAGGTCTG 47
 |||||||
 Db 386 GRGCTGCCGAGGTCTG 370

US-09-161-092-3/c
 Sequence 3, Application US/09161092
 ; Parent No. 6301314
 ; GENERAL INFORMATION:
 ; APPLICANT: ALLAN, Gordon
 ; APPLICANT: MEEHAN, Brian
 ; APPLICANT: CLARK, Edward
 ; APPLICANT: HAINES, Deborah
 ; APPLICANT: HASSARD, Lori
 ; APPLICANT: HARDING, John
 ; APPLICANT: CHARREYRE, Catherine E.
 ; APPLICANT: CHAPPUIS, Gilles E.
 ; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
 ; TITLE OF INVENTION: REAGENTS
 ; FILE REFERENCE: ALLAN
 ; CURRENT APPLICATION NUMBER: US/09/161,092
 ; CURRENT FILING DATE: 1998-09-25
 ; PRIORITY NUMBER: EARLIER APPLICATION NUMBER: 09/082,558
 ; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-05-21
 ; PRIORITY NUMBER: EARLIER APPLICATION NUMBER: FR 9800873
 ; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-01-22
 ; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9803707
 ; PRIORITY FILING DATE: EARLIER FILING DATE: 1998-03-20
 ; PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 97/12382
 ; PRIORITY FILING DATE: EARLIER FILING DATE: 1997-10-03
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: Patentin Ver. 2.0
 ; SEQ ID NO 3
 ; LENGTH: 1768

Query Match 3.6%; Score 17; DB 4; Length 1768;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 SEQ ID NO 3
 Qy 31 GTGCTGCCGAGGTCTG 47
 |||||||
 Db 385 GTGCTGCCGAGGTCTG 369

RESULT 12
 US-09-082-558-6/c

```

; TYPE: DNA
; ORGANISM: Porcine circovirus
; US-09-161-092-3

Query Match          3.6%; Score 17; DB 4; Length 1768;
Best Local Similarity 100.0%; Pred. No. 29; Matches 0; Indels 0; Gaps 0;
QY      31 GTGCTGCCAGGTGCTG 47
Db      385 GTGCTGCCAGGTGCTG 369

RESULT 14
; Sequence 4, Application US/09161092
; Patent No. 6391314
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEEHAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HASSARD, Lori
; APPLICANT: HARDING, John
; APPLICANT: CHARREIRE, Catherine E.
; APPLICANT: CHAPPOIS, Gilles E.
; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/161.092
; CURRENT FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/082, 558
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9800873
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9803707
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 97/12382.
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-03
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 1768
; TYPE: DNA
; ORGANISM: Porcine circovirus
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)..(1768)
; OTHER INFORMATION: N represents A or C or G or T
; US-09-161-092-6

Query Match          3.6%; Score 17; DB 4; Length 1768;
Best Local Similarity 100.0%; Pred. No. 29; Matches 0; Indels 0; Gaps 0;
QY      31 GTGCTGCCAGGTGCTG 47
Db      386 GTGCTGCCAGGTGCTG 370

Search completed: July 20, 2003, 07:21:09
Job time : 45 secs

RESULT 15
; Sequence 6, Application US/09161092
; Patent No. 6391314
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEEHAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HASSARD, Lori
; APPLICANT: HARDING, John
; APPLICANT: CHARREIRE, Catherine E.
; APPLICANT: CHAPPOIS, Gilles E.
; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/161.092
; CURRENT FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/082, 558
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-05-21
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9800873
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 9803707
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: FR 97/12382.
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-03
; PRIOR FILING DATE: EARLIER FILING DATE: 1997-10-03
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO: 6
; LENGTH: 1768
; TYPE: DNA
; ORGANISM: Porcine circovirus
; US-09-161-092-4

Query Match          3.6%; Score 17; DB 4; Length 1768;
Best Local Similarity 100.0%; Pred. No. 29; Matches 0; Indels 0; Gaps 0;
QY      31 GTGCTGCCAGGTGCTG 47
Db      385 GTGCTGCCAGGTGCTG 369

RESULT 15
; Sequence 6, Application US/09161092
; Patent No. 6391314
; GENERAL INFORMATION:
; APPLICANT: ALLAN, Gordon
; APPLICANT: MEEHAN, Brian
; APPLICANT: CLARK, Edward
; APPLICANT: HAINES, Deborah
; APPLICANT: HASSARD, Lori
; APPLICANT: HARDING, John
; APPLICANT: CHARREIRE, Catherine E.
; APPLICANT: CHAPPOIS, Gilles E.
; TITLE OF INVENTION: NEW PORCINE CIRCOVIRUSES, VACCINES AND DIAGNOSTIC
; TITLE OF INVENTION: REAGENTS
; FILE REFERENCE: ALLAN
; CURRENT APPLICATION NUMBER: US/09/161.092

```

ON nucleic - nucleic search, using sw model.

Run on: July 20, 2003, 06:51:05 ; Search time 172 Seconds
 (without alignments)
 5697.240 Million cell updates/sec

Title: US-09-719-002-1

Perfect score: 475

Sequence: 1 gaattttatgtggactga.....acaacacacaaacatcg 475

Scoring table: ORIGO.NUC

Gapop 60.0 , Gapext 60.0

Total number of hits satisfying chosen parameters: 2879534

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Published Applications_NA:*

```

1: /cgn2_6/ptodata/1/pubpna/us07_pubcomb.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/us06_pubcomb.seq:*
4: /cgn2_6/ptodata/1/pubpna/us07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/1/pubpna/PCITS_PUBCOMB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCITS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/us08_pubcomb.seq:*
8: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/us09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/1/pubpna/us09_NEW_PUB.seq:1:*
11: /cgn2_6/ptodata/1/pubpna/us09_NEW_PUB.seq:2:*
12: /cgn2_6/ptodata/1/pubpna/us09_NEW_PUB.seq:3:*
13: /cgn2_6/ptodata/1/pubpna/us09_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/us10_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/us10_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/us60_NEW_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/us60_PUBCOMB.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result NO.	Score	Query Match Length	DB ID	Description
1	21	4.4	716	US-10-027-632-34117 Sequence 34117, A
2	21	4.4	716	US-10-027-632-34118 Sequence 34118, A
c - 4	21	4.4	1691139	US-10-027-510-1 Sequence 1, Appli
c - 4	19	4.0	584	US-10-027-632-243175 Sequence 243175,
c - 5	19	4.0	649	US-10-027-632-134120 Sequence 134120,
c - 6	19	4.0	649	US-10-027-632-134121 Sequence 134121,
c - 7	18	3.8	397	US-09-917-800-897 Sequence 897, APP
c - 8	18	3.8	475	US-09-918-995-2571 Sequence 25871, A
c - 9	18	3.8	491	US-10-027-632-48705 Sequence 48705, A
c - 10	18	3.8	491	US-10-027-632-48705 Sequence 48706, A
c - 11	18	3.8	612	US-10-027-632-61505 Sequence 61505, A
c - 12	18	3.8	612	US-10-027-632-62397 Sequence 62397, A
c - 13	18	3.8	612	US-10-027-632-63229 Sequence 63229, A
c - 14	18	3.8	656	US-10-027-632-11356 Sequence 111366, A
c - 15	18	3.8	694	US-10-027-632-48227 Sequence 48227, A
c - 16	18	3.8	745	US-10-027-632-28109 Sequence 28109, A

ALIGNMENTS

RESULT 1
 Sequence 34117, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 TITLE OF INVENTION: Polymorphisms in the Human Genome
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10-027, 632
 CURRENT FILING DATE: 2002-04-30
 PRIOR APPLICATION NUMBER: US 60/218, 006
 PRIOR FILING DATE: 2000-07-12
 PRIOR APPLICATION NUMBER: US 60/198, 676
 PRIOR FILING DATE: 2000-04-20
 PRIOR APPLICATION NUMBER: US 60/193, 483
 PRIOR FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: US 60/185, 218
 PRIOR FILING DATE: 2000-02-24
 PRIOR APPLICATION NUMBER: US 60/167, 363
 PRIOR FILING DATE: 1999-11-23
 PRIOR APPLICATION NUMBER: US 60/156, 358
 PRIOR FILING DATE: 1999-09-28
 PRIOR APPLICATION NUMBER: US 60/146, 002
 PRIOR FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 34117
 LENGTH: 716
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-34117

Query Match Best Local Similarity 100.0%; Pred. No. 0.86; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 241 CTACATTAATTCACATTG 261

Db 361 CTACATTAATTCACATTG 381

RESULT 2-632-34118
 US-10-027-632-34118
 ; Sequence 34118, Application US/10027632
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 34118
 LENGTH: 716
 ; TYPE: DNA
 ; ORGANISM: Human
 ;
 US-10-027-632-34118
 ; Query Match 4.4%; Score 21; DB 15; Length 716;
 ; Best Local Similarity 100.0%; Pred. No. 0.86; 0; Mismatches 0; Indels 0; Gaps 0;
 ; Matches 21; Conservative 0; ;
 Qy 241 CTACATTAACTACACTTGTG 261
 Db 361 CTACATTAACTACACTTGTG 381
 ;
 RESULT 3
 US-10-067-514-1
 ; Sequence 1, Application US/10067514
 ; GENERAL INFORMATION:
 ; APPLICANT: Grethrðsdóttir, Sölveig
 ; APPLICANT: Jónsdóttir, Sif
 ; APPLICANT: Reynisdóttir, Sigríður Th.
 ; TITLE OF INVENTION: HUMAN STROKE GENE
 ; FILE REFERENCE: 2345-2010-003
 ; CURRENT APPLICATION NUMBER: US/10/067,514
 ; CURRENT FILING DATE: 2002-02-04
 ; PRIOR APPLICATION NUMBER: US 09/811/352
 ; PRIOR FILING DATE: 2001-03-19
 ; NUMBER OF SEQ ID NOS: 84
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 1
 LENGTH: 1691139
 ; TYPE: DNA
 ; ORGANISM: Human
 ;
 US-10-067-514-1
 ; Query Match 4.4%; Score 21; DB 15; Length 1691139;
 ; Best Local Similarity 100.0%; Pred. No. 1.4; 0; Mismatches 0; Indels 0; Gaps 0;
 ; Matches 21; Conservative 0; ;
 Qy 241 CTACATTAACTACACTTGTG 261
 Db 761375 CTACATTAACTACACTTGTG 761395
 ;
 RESULT 4
 US-10-027-632-243175/C
 ; Sequence 243175, Application US/10027632
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 23175
 LENGTH: 584
 ; TYPE: DNA
 ; ORGANISM: Human
 ; FEATURE: misc_feature
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)-(584)
 ; OTHER INFORMATION: n = A,T,C or G
 ;
 US-10-027-632-243175
 ; Query Match 4.0%; Score 19; DB 15; Length 584;
 ; Best Local Similarity 100.0%; Pred. No. 9.7; 0; Mismatches 0; Indels 0; Gaps 0;
 ; Matches 19; Conservative 0; ;
 Qy 254 AACTTTGCAATAAATA 272
 Db 156 AACTTTGCAATAAATA 138
 ;
 RESULT 5
 US-10-027-632-134120
 ; Sequence 134120, Application US/10027632
 ; GENERAL INFORMATION:
 ; APPLICANT: Wang, David G.
 ; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 ; FILE REFERENCE: 108827.129
 ; CURRENT APPLICATION NUMBER: US/10/027,632
 ; CURRENT FILING DATE: 2002-04-30
 ; PRIOR APPLICATION NUMBER: US 60/218,006
 ; PRIOR FILING DATE: 2000-07-12
 ; PRIOR APPLICATION NUMBER: US 60/198,676
 ; PRIOR FILING DATE: 2000-04-20
 ; PRIOR APPLICATION NUMBER: US 60/193,483
 ; PRIOR FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: US 60/185,218
 ; PRIOR FILING DATE: 2000-02-24
 ; PRIOR APPLICATION NUMBER: US 60/167,363
 ; PRIOR FILING DATE: 1999-11-23
 ; PRIOR APPLICATION NUMBER: US 60/156,358
 ; PRIOR FILING DATE: 1999-09-28
 ; PRIOR APPLICATION NUMBER: US 60/146,002
 ; PRIOR FILING DATE: 1999-08-09
 ; NUMBER OF SEQ ID NOS: 325720
 ; SOFTWARE: FastSEQ for Windows Version 4.0
 ; SEQ ID NO 134120

```

; LENGTH: 649
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-134120

Query Match          4.0%; Score 19; DB 15; Length 649;
Best Local Similarity 100.0%; Pred. No. 9.8; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      418 CTGCCATTCCCTCTCC 436
Db      433 CTGCCATTCCCTCTCC 451

RESULT 6           US-10-027-632-134121
; Sequence 154121; Application US/10027632
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108427-129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO: 134121
; LENGTH: 649
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-134121

Query Match          4.0%; Score 19; DB 15; Length 649;
Best Local Similarity 100.0%; Pred. No. 9.8; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      418 CTGCCATTCCCTCTCC 436
Db      433 CTGCCATTCCCTCTCC 451

RESULT 7           US-09-917-800A-897
; Sequence 897; Application US/09917800A
; Patent No. US20020119462A1
; GENERAL INFORMATION:
; APPLICANT: Mendrick, Donna
; APPLICANT: Porter, Mark
; APPLICANT: Johnson, Kory
; APPLICANT: Castle, Arthur
; APPLICANT: Elashoff, Michael
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Molecular Toxicology Modeling
; FILE REFERENCE: 44931-5038-US
; CURRENT APPLICATION NUMBER: US/09/917,800A
; CURRENT FILING DATE: 2001-07-31
; PRIOR APPLICATION NUMBER: US 60/222,040
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: US 60/222,880

Query Match          4.0%; Score 19; DB 15; Length 649;
Best Local Similarity 100.0%; Pred. No. 9.8; Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      418 CTGCCATTCCCTCTCC 436
Db      433 CTGCCATTCCCTCTCC 451

RESULT 8           US-09-918-995-25971/c
; Sequence 25971; Application US/09918995
; Publication No. US20030073523A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED FROM VARIOUS CDNA LIBRARIES
; FILE REFERENCE: 2011-756
; CURRENT APPLICATION NUMBER: US/09/918,995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: Fast-SEQ for Windows Version 3.0
; SEQ ID NO: 25871
; LENGTH: 475
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1..(475))
; OTHER INFORMATION: n = A,T,C or G
; US-09-918-995-25971

Query Match          3.8%; Score 18; DB 12; Length 475;
Best Local Similarity 100.0%; Pred. No. 32; Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      263 AAATAAATATAAGAAA 280
Db      68 AAATAAATATAAGAAA 51

```

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIORITY APPLICATION NUMBER: US 60/218,006
 PRIORITY FILING DATE: 2000-07-12
 PRIORITY APPLICATION NUMBER: US 60/198,676
 PRIORITY FILING DATE: 2000-04-20
 PRIORITY APPLICATION NUMBER: US 60/193,483
 PRIORITY FILING DATE: 2000-03-29
 PRIORITY APPLICATION NUMBER: US 60/185,218
 PRIORITY FILING DATE: 2000-02-24
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-11-23
 PRIORITY APPLICATION NUMBER: US 60/156,358
 PRIORITY FILING DATE: 1999-09-28
 PRIORITY APPLICATION NUMBER: US 60/146,002
 PRIORITY FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 48705
 LENGTH: 491
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-48706
 Sequence 61505, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIORITY APPLICATION NUMBER: US 60/218,006
 PRIORITY FILING DATE: 2000-07-12
 PRIORITY APPLICATION NUMBER: US 60/198,676
 PRIORITY FILING DATE: 2000-04-20
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-11-23
 PRIORITY APPLICATION NUMBER: US 60/156,358
 PRIORITY FILING DATE: 1999-09-28
 PRIORITY APPLICATION NUMBER: US 60/146,002
 PRIORITY FILING DATE: 2000-02-24
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 61505
 LENGTH: 512
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-61505
 Query Match 3.8%; Score 18; DB 15; Length 491;
 Best Local Similarity 100.0%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Dd
 Db
 439 CAAATAATTATTCAGT 456
 RESULT 10
 US-10-027-632-48706
 Sequence 48706, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIORITY APPLICATION NUMBER: US 60/218,006
 PRIORITY FILING DATE: 2000-07-12
 PRIORITY APPLICATION NUMBER: US 60/198,676
 PRIORITY FILING DATE: 2000-04-20
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-11-23
 PRIORITY APPLICATION NUMBER: US 60/156,358
 PRIORITY FILING DATE: 1999-09-28
 PRIORITY APPLICATION NUMBER: US 60/146,002
 PRIORITY FILING DATE: 2000-02-24
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 48706
 LENGTH: 491
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-48706
 Query Match 3.8%; Score 18; DB 15; Length 491;
 Best Local Similarity 100.0%; Pred. No. 32; Mismatches 0; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Dd
 Db
 100 CAAATAATTATTCAGT 117
 RESULT 11
 US-10-027-632-61505
 Sequence 61505, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIORITY APPLICATION NUMBER: US 60/218,006
 PRIORITY FILING DATE: 2000-07-12
 PRIORITY APPLICATION NUMBER: US 60/198,676
 PRIORITY FILING DATE: 2000-04-20
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-11-23
 PRIORITY APPLICATION NUMBER: US 60/156,358
 PRIORITY FILING DATE: 1999-09-28
 PRIORITY APPLICATION NUMBER: US 60/146,002
 PRIORITY FILING DATE: 2000-02-24
 PRIORITY APPLICATION NUMBER: US 60/167,363
 PRIORITY FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 61505
 LENGTH: 612
 TYPE: DNA
 ORGANISM: Human
 US-10-027-632-61505
 Query Match 3.8%; Score 18; DB 15; Length 612;
 Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0; Indels 0; Gaps 0;
 Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Dd
 Db
 402 AATAAAATATAAGAAA 419
 RESULT 12
 US-10-027-632-62397
 Sequence 62397, Application US/10027632
 GENERAL INFORMATION:
 APPLICANT: Wang, David G.
 TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
 FILE REFERENCE: 108827-129
 CURRENT APPLICATION NUMBER: US/10/027,632
 CURRENT FILING DATE: 2002-04-30
 PRIORITY APPLICATION NUMBER: US 60/218,006
 PRIORITY FILING DATE: 2000-07-12
 PRIORITY APPLICATION NUMBER: US 60/198,676
 PRIORITY FILING DATE: 2000-04-20
 PRIORITY APPLICATION NUMBER: US 60/193,483
 PRIORITY FILING DATE: 2000-03-29
 PRIORITY APPLICATION NUMBER: US 60/156,358
 PRIORITY FILING DATE: 1999-09-28
 PRIORITY APPLICATION NUMBER: US 60/146,002
 PRIORITY FILING DATE: 1999-08-09
 NUMBER OF SEQ ID NOS: 325720
 SOFTWARE: FastSEQ for Windows Version 4.0
 SEQ ID NO: 62397
 LENGTH: 612

```

; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-62397
;
; Query Match_h          3.8%; Score 18; DB 15; length 612;
; Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0;
; Matches 18; Conservative 0; Indels 0; Gaps 0;
;
; Qy      263 AAATAAATATAAGAAA 280
; Db      ||||||| 402 AAATAAATATAAGAAA 419
;

RESULT 13
US-10-027-632-63229
; Sequence 63229; Application US/10027632
;
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827-129
;
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
;
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
;
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1998-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 63229
; LENGTH: 612
;
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-63229

Query Match_h          3.8%; Score 18; DB 15; length 656;
; Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0;
; Matches 18; Conservative 0; Indels 0; Gaps 0;
;
; Qy      444 TCTACTCAAAACACA 461
; Db      385 TCTACTCAAAACACA 368
;

RESULT 15
US-10-027-632-48227
; Sequence 48227; Application US/10027632
;
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827-129
;
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
;
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 48227
; LENGTH: 694
;
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-48227

Query Match_h          3.8%; Score 18; DB 15; length 612;
; Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0;
; Matches 18; Conservative 0; Indels 0; Gaps 0;
;
; Qy      263 AAATAAATATAAGAAA 280
; Db      ||||||| 402 AAATAAATATAAGAAA 419
;

RESULT 14
US-10-027-632-111366c
; Sequence 111366; Application US/10027632
;
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827-129
;
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
;
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO: 48227
; LENGTH: 694
;
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-48227

Query Match_h          3.8%; Score 18; DB 15; length 694;
; Best Local Similarity 100.0%; Pred. No. 33; Mismatches 0;
; Matches 18; Conservative 0; Indels 0; Gaps 0;
;
; Qy      263 AAATAAATATAAGAAA 280
; Db      392 AAATAAATATAAGAAA 409
;

Search completed: July 20, 2003, 07:53:40
Job time : 176 secs

```



OK nucleic - nucleic search, using sw model
 Run on: July 20, 2003, 06:08:35 ; Search time 1127 Seconds
 (without alignments)
 6825.959 Million cell updates/sec

Title: US-09-719-002-1
 Perfect score: 475
 Sequence: 1 gaattcttatttgcgactga.....acaacacacaacaaatcatg 475
 Scoring table: OLIGO_NUC
 Gappp 60.0 , Gapext 60.0

Searched: 16154066 seqs, 8097743376 residues
 Word size : 0
 Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
 Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries
 Database : EST:*

1: em_estba:*

2: em_lesthum:*

3: em_lestin:*

4: em_estmu:*

5: em_lestov:*

6: em_lestapl:*

7: em_lestro:*

8: em_hcc:*

9: qb_est1:*

10: qb_est2:*

11: qb_htc:*

12: qb_est3:*

13: qb_est4:*

14: qb_est5:*

15: em_estfun:*

16: em_estom:*

17: qb_gss:*

18: em_gss_hum:*

19: em_gss_inv:*

20: em_gss_pIn:*

21: em_gss_vrt:*

22: em_gss_fun:*

23: em_gss_mam:*

24: em_gss_mus:*

25: em_gss_other:*

26: em_gss_pro:*

27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

-result	%	Query Match Length	DB ID	Description
1	21	4.4	252	10 BE370506
2	21	4.4	426	10 BE147853
C	3	4.4	587	10 BE205326
C	4	4.4	669	13 BI311957
5	21	4.4	769	17 AG143540
6	21	4.4	822	17 BH371846

SUMMARIES

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1 (bases 1 to 25 (25))	NIR-MGC	http://mgc.ncbi.nih.gov/	National Institutes of Health, Mammalian Gene Collection (MGC)	Contact: Robert Strausberg, Ph.D. Email: cgbps1@mail.nih.gov

FEATURES

Source
1 .232

ALIGNMENTS

RESULT 1
 BE370506 LOCUS BB370506 DEFINITION 601218294F1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:3187470 5', mRNA sequence.
 ACCESSION BB370506 VERSION BB370506.1 GI:9315869 KEYWORDS EST.* SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; vertebrata; Euteleostomi; Mammalia; Buteraria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 1 (bases 1 to 25 (25))
 NIH-MGC http://mgc.ncbi.nih.gov/
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:
 http://Image.lnl.gov
 Plate: LIAM8750 row: k column: 23
 High quality sequence stop: 242.
 Location/Qualifiers

BASE COUNT	109 a	105 c	121 g	91 t	ORIGIN
Query Match	4.4%	Score 21;	DB 10;	Length 587;	
Best Local Similarity	100.0%	Pred. No. 5;	Mismatches	0;	
Matches	21;	Conservative	0;	Indels	0;
QY	449	CTCAAAACACACACACCA	469	Gaps	0;
Db	406	CTCAAAACACACACACACCA	426		
BASE COUNT	75 a	61 c	43 g	73 t	ORIGIN
Query Match	4.4%	Score 21;	DB 10;	Length 426;	
Best Local Similarity	100.0%	Pred. No. 4.5;	Mismatches	0;	
Matches	21;	Conservative	0;	Indels	0;
QY	454	AAACACACACACACAT	474	Gaps	0;
Db	54	AAACACACACACACAT	34		
RESULT 2					
BEI47853					
LOCUS	BEI47853	426 bp	mRNA	linear	EST 21-JUN-2000
DEFINITION	RC3-HT0230-201199-013-g05	HT0230	Homo sapiens	cDNA,	mRNA sequence.
ACCESSION	BEI47853				
VERSION	275				
KEYWORDS					
SOURCE	232 ACTTGTGCAATAATATAA	252			
ORGANISM					
REFERENCE					
AUTHORS					
1 (bases 1 to 426)					
Mammalia; Eutheria; Primates; Catarrhini; Homidae; HOMO.					
Nogai, M.A., da Silva, C., Correa, R., Verjovski-Almeida, S., Briones, M.R.,					
Goldman, G.H., Carvalho, A.P., Bordim, S., Costa, F.F., Costa, F.F.,					
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare					
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and					
Simpson, A.J.					
TITLE					
JOURNAL					
MEDLINE					
COMMENT					
FEATURES					
Source					
RESULT 3					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
Source					
RESULT 4					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 5					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 6					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 7					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 8					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 9					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM					
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
Spermatophyta; Magnoliophyta; eudicots; core eudicots;					
Rosidae; eurosids I; Fabales; Papilionoideae; Trifolieae;					
Medicago; eurosid I; Fabaceae; Papilionoideae; Trifolieae;					
1 (bases 1 to 587)					
VandenBosch, K., Endre, G., Hur, J., Moore, J., Beremand, P., Ellis, L.,					
Town, C.D., Bowmen, C.L., Craven, M.B., Hansen, T.S., Holt, J.E. and					
Fraser, C.M.					
BB20 Biocenter, 1445 Gortner Ave, St. Paul, MN 55108, USA					
Tel: 612 624 2755					
Fax: 612 625 1738					
Email: kvandenbosch.umn.edu					
TIGR sequence name: T665098E					
More information is available at:					
http://chrysise.tamu.edu/medicago/					
Seq primer: Skmod (CTA GAA CTA gtg GAT CC).					
1. 587					
FEATURES					
source					
RESULT 10					
BB205326/c					
LOCUS	BB205326	587 bp	mRNA	linear	EST 05-SEP-2000
DEFINITION	EST398002	KV0	Medicago truncatula	cDNA clone	PKV0-21E10, mRNA
ACCESSION	BB205326				
VERSION	1				
KEYWORDS					
SOURCE					
ORGANISM	"Mus musculus"				
/strain	"C57BL/10ScN"				
/db_xref	"txon:1009"				
/clone	"IMAGE:3587470"				
/tissue	"spontaneous tumor, metastatic to mammary."				
/lab_host	"DH10B"				
/note	"Organ: lung; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator: Gilbert Smith, NIH"				
Db					
BASE COUNT	164 a	69 c	143 g	211 t	ORIGIN
Query Match	4.4%	Score 21;	DB 10;	Length 587;	
Best Local Similarity	100.0%	Pred. No. 5;	Mismatches	0;	
Matches	21;	Conservative	0;	Indels	0

RESULT 4
31311057/c
LOCUS B1311957 659 bp mRNA linear EST 20-JUL-2001
DEFINITION EST531370. GSSD Medicago truncatula cDNA clone PGESD16D21 5' end,
mRNA sequence.
ACCESSION B1311957
VERSION B1311957.1 GI:14986284
KEYWORDS EST
SOURCE barrel medic.

REFERENCE
AUTHORS Grusak,M.A., Samac,D.A., Town,C.D., Van Aken,S., Utterback,T., Cho,J., and Fraser,C.M.
TITLE Unpublished (2001)
JOURNAL Contact: Michael A. Grusak
COMMENT USDA/FARS Children's Nutrition Research Center
Baylor College of Medicine
1100 Bates Street, Houston, TX 77030-2600, USA
Tel: 713-798-7044
Fax: 713-798-7048
Email: mgrusak@bcm.tmc.edu
B398812e

More information is available at: www.medicago.org
Seq primer: SKmod (CTA GAA CTA gtg gAT CC).
FEATURES source

```

<!ELEMENT source (clone_id, organism, tissue_type, dev_stage, pollination, note, xref, db_xref, clone_lib, db, cultivar, sex)>
<!ELEMENT clone_id ("PGESD16D21")>
<!ELEMENT organism ("Medicago truncatula")>
<!ELEMENT tissue_type ("immature seeds")>
<!ELEMENT dev_stage ("Immature seeds, 11 to 19 days after pollination")>
<!ELEMENT pollination ("Vector: pBlue-script SK-, Site:1: EcoRI; Site:2: XbaI; Immature seeds, collected from pods ranging in age from 11 to 19 days after pollination, were harvested from greenhouse-grown plants. Seeds were removed and separated from pod walls and were immediately frozen in liquid nitrogen. Seeds throughout the age range were pooled for mRNA extraction. cDNA was prepared from polyA+ enriched RNA. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in XLOR cells.">
<!ELEMENT note ("The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-zap phage using Ex-assist helper phage and propagated in XLOR cells.")>
<!ELEMENT xref ("PRC:380")>
<!ELEMENT db_xref ("PGESD16D21")>
<!ELEMENT clone_lib ("GSSD")>
<!ELEMENT db ("PGESD16D21")>
<!ELEMENT cultivar ("AI")>
<!ELEMENT sex ("Male")>
<!ELEMENT cell_type ("lymphocytes")>
<!ELEMENT clone_id ("RPCI-43 Chimpanzee Male BAC Library")>
<!ELEMENT clone_lib ("RPCI-43 Chimpanzee Male BAC Library")>
```

BASE COUNT 178 a 186 a 203 c 109 g 270 t 2 others
ORIGIN 86 c 182 g 223 t

RESULT 5
AC143540 LOCUS AG143540 769 bp DNA linear GSS 08-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-004E06.TU, genomic survey
SEQUENCE
-D- 571 AACACACACACACCATCAT 551
-D- 454 AACACACACACACCATCAT 474
QY

RESULT 5
AC143540 LOCUS AG143540 769 bp DNA linear GSS 08-JAN-2002
DEFINITION Pan troglodytes DNA, clone: RP43-004E06.TU, genomic survey
SEQUENCE
ACCESSION AG143540

VERSION AG143540.1 GI:16673218
KEYWORDS GSS.
SOURCE Pan troglodytes male lymphocytes DNA, clone_lib:RPCI-43 Chimpanzee
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. Male BAC Library clone:RP43-004E05.TU.

REFERENCE
AUTHORS Fujimura,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Totoki,Y., Watenabe,H. and Sasaki,Y.
TITLE JOURNAL Unpublished (bases 1 to 769)
COMMENT Submitted (02-AUG-2001) Asao Fujimura, The Institute of Physical and Chemical Research (Riken), Genomic Sciences Center (GSC), 1-7-22 Suenriro-chou,Tsurumi-ku, Yokohama, Kanagawa 230-0005, Japan (E-mail:chimpanzee@gscreiken.go.jp, URL:<http://hgp.gsc.riken.go.jp/>, Tel:81-45-503-9111, Fax:81-45-503-9170). Clones are derived from the chimpanzee BAC library RPCI-43 This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
PRIMERS Sequencing: TJ
LIBRARY Vector : PBACE3.6
FEATURES source
R.SITE 1 : ECORI
R.SITE 2 : ECORI
LOCATION/QUALIFIERS 1. .769
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-004E06.TU"
/sex="male"
/cell_type="lymphocytes"
/clone_id="RPCI-43 Chimpanzee Male BAC Library"
/clone_lib="RPCI-43 Chimpanzee Male BAC Library"
BASE COUNT 185 a 203 c 109 g 270 t 2 others
ORIGIN 86 c 182 g 223 t

RESULT 6
BH371846 LOCUS BH371846 822 bp DNA linear GSS 10-DEC-2001
DEFINITION AG-ND-162M17.TF ND-TAM Anopheles gambiae genomic clone AG-ND-162M17
ACCESSION BH371846
VERSION BH371846.1 GI:17317971
KEYWORDS GSS.
SOURCE African malaria mosquito.
ORGANISM Anopheles gambiae
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anophelidae.
REFERENCE
AUTHORS Shetty,J., Malek,J., Koo,H., Collins,F., Gardner,M. and Loftus,B.J.
TITLE JOURNAL Unpublished (2001)
COMMENT Other_GSS: AG-ND-162M17.TF
Contact: Brendan J Loftus
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 88 0208

Fax: 301 838 3543

Email: bjloufus@tigr.org

This clone is from an *A. gambiae* BAC library (ND-TAM) provided by F.H. Collins and sequenced by The Institute for Genomic Research (TIGR). The BAC library was generated from *A. gambiae* PEST Strain DNA. All DNA was extracted from newly hatched first instar larvae to minimize the inclusion of DNA from microorganisms that inhabit the gut. The DNA is derived from mixed sexes of larvae. The BAC library was constructed at Texas A&M University BAC Center University, College Station, Texas 77843-2123, USA using a HindIII partial digest.

Seq Primer: M13 For

Class: BAC ends.

FEATURES

source

BASE COUNT	ORIGIN	QY	DB
235 a	142 c	337	TTCACACAAGGCTAGACTT 357
235 a	155 g	394	TTCACACAGGCTTAGACTT 414
RESULT 7	RESULT 8	RESULT 9	RESULT 10
BF573809	BF573809	CNS030N2	CNS030N2
DEFINITION mRNA sequence.	DEFINITION mRNA sequence.	LOCUS TETRAODON nigroviridis	LOCUS TETRAODON nigroviridis
ACCESSION BF573809	ACCESSION BF573809	DEFINITION GSS: genome survey sequence.	DEFINITION GSS: genome survey sequence.
VERSION 2.1	VERSION 2.1	VERSION A12245.5	VERSION A12245.5
KEYWORDS EST;	KEYWORDS EST;	KEYWORDS Tetraodon nigroviridis	KEYWORDS Tetraodon nigroviridis
SOURCE human.	SOURCE human.	SOURCE Organism	SOURCE Organism
ORGANISM Homo sapiens	ORGANISM Homo sapiens	ORGANISM Actinopterygii	ORGANISM Actinopterygii
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Bony fish; Teleostei; Neoteleostei; Tetraodontiformes; Tetraodontidae; Tetraodon.	Mammalia; Eutheria; Primates; Catarrini; Homidae; Homo. I (bases 1 to 845) NIH MGC http://mgc.nci.nih.gov/.	Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi; Bony fish; Teleostei; Neoteleostei; Tetraodontiformes; Tetraodontidae; Tetraodon.	Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Bony fish; Teleostei; Neoteleostei; Tetraodontiformes; Tetraodontidae; Tetraodon.
REFERENCE Bernot,A., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Bounau,L., Fisher,C., Bernot,A., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Saurin,W., and Weissenbach,J.	REFERENCE Bernot,A., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Saurin,W., and Weissenbach,J.	REFERENCE Bernot,A., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Saurin,W., and Weissenbach,J.	REFERENCE Bernot,A., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Roest-Crollius,H., Jaillon,O., Basilia,C., Flizanes,C., Fisher,C., Saurin,W., and Weissenbach,J.
AUTHORS	AUTHORS	AUTHORS	AUTHORS
JOURNAL Unpublished (1999); Contact: Robert Strausberg, Ph.D.	JOURNAL Unpublished (1999); Contact: Robert Strausberg, Ph.D.	JOURNAL Unpublished (1999); Contact: Robert Strausberg, Ph.D.	JOURNAL Unpublished (1999); Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgabbs-r@mail.nih.gov	COMMENT Email: cgabbs-r@mail.nih.gov	COMMENT Email: cgabbs-r@mail.nih.gov	COMMENT Email: cgabbs-r@mail.nih.gov
Tissue Procurement: CLONETECH Laboratories, Inc.	Tissue Procurement: CLONETECH Laboratories, Inc.	TITLE Direct Submission	TITLE Direct Submission
CDNA Library Preparation: CLONETECH Laboratories, Inc.	CDNA Library Preparation: CLONETECH Laboratories, Inc.	REFERENCE 1 (bases 1 to 970)	REFERENCE 1 (bases 1 to 970)
CDNA Library Arrayed by: The T.M.A.G.E. Consortium (LLNL)	CDNA Library Arrayed by: The T.M.A.G.E. Consortium (LLNL)	AUTHORS Bonnaud,L., Billault,A., Quetier,F., Saurin,W., Bernot,A., and Weissenbach,J.	AUTHORS Bonnaud,L., Billault,A., Quetier,F., Saurin,W., Bernot,A., and Weissenbach,J.
DNA Sequencing by: Invitae Genomics, Inc.	DNA Sequencing by: Invitae Genomics, Inc.	JOURNAL Unpublished (12-APR-2000)	JOURNAL Unpublished (12-APR-2000)
Clone distribution: MGC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at:	Clone distribution: MGC clone distribution information can be found through the T.M.A.G.E. Consortium/LLNL at:	CHARACTERIZATION AND REPEAT ANALYSIS OF THE COMPACT GENOME OF THE FRESHWATER PUFFERFISH TETRAODON NIGROVIRIDIS	CHARACTERIZATION AND REPEAT ANALYSIS OF THE COMPACT GENOME OF THE FRESHWATER PUFFERFISH TETRAODON NIGROVIRIDIS
Plate: IICM1083 Row: 1 column: 01	Plate: IICM1083 Row: 1 column: 01	COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cnrs.fr/~tetraodon/ .	COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cnrs.fr/~tetraodon/ .
FEATURES source	FEATURES source	FEATURES source	FEATURES source
1. .845 /organism="Homo sapiens"	1. .970 /organism="Tetraodon nigroviridis"	1. .970 /organism="Tetraodon nigroviridis"	1. .970 /organism="Tetraodon nigroviridis"
/db_xref="taxon:9606"	/db_xref="taxon:9983"	/db_xref="taxon:9983"	/db_xref="taxon:9983"
/clone="IMAGE:4271376"	/clone="203E01"	/clone="203E01"	/clone="203E01"
/clone_id="NTB_MG_81"	/clone_id="G"	/clone_id="G"	/clone_id="G"
/lab_host="DHLOB (T1 phage-resistant)" /notes="Organ: muscle (skeletal); vector: pDNR-LIB (Clonetech); Site_1: SfiI (ggccggctcgcc); Site_2: SfiI (ggccattatggcc); 5' and 3' adaptors were used in cloning, as follows: 5' adaptor sequence: 5'-CACGCCATATTGGCC-3' and 3' adaptor sequence: 5'-ATTCCTGAGGGCGAGCGACAGC-3' (where B = A,	/notes="Genoscope sequence ID : COAS203ac01P1-end : T7"	/notes="Genoscope sequence ID : COAS203ac01P1-end : T7"	/notes="Genoscope sequence ID : COAS203ac01P1-end : T7"
BASE COUNT	ORIGIN	QY	DB
133 a	334 c	32	TGCTGCCGAGGTGTTGCAA 52
342 g	145 t	480	TGCTGCCGAGGTGTTGCAA 500

C, or G and N = A, C, G, or T). Average insert size 1.55 kb (range 1.0-4.0 kb).

15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA).

BASE COUNT	127	ORIGIN	a
RESULT	11	Query Match	4.2%
LOCUS	AA461949	Score	20;
DEFINITION	v971c09_r1 Soares mouse	DB	13;
REFERENCE	NM_0010005	Length	401;
AUTHORS	M., mRNA sequence	EST	10-JUN-1996
ACCESSION	AAB461949	bp	
VERSION	AA461949.1	mRNA	
KEYWORDS	EST	linear	
SOURCE	EST	EST	
ORGANISM	Mus musculus	cdNA clone	
	House mouse.	IMAGE:871408	
	Mus musculus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus		
	1 (bases 1 to 405)		
	María M., Hillier,T., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thaising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.		
TITLE	The WashU-HMM Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: María M./Mouse EST Project		
	WashU-HMM Mouse EST Project		
	Washington University School of Medicine		
	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108		
	Tel: 314 286 1800		
	Fax: 314 286 1810		
	Email: mouseest@atson.wustl.edu		
	This clone is available royalty-free through LINL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.		
	NCI:51088		
	Seq primer: -28m13 rev2 ET from Amersham		
	High quality sequence stop: 353.		
FEATURES	Location/Qualifiers		
source			
	1. . 405		
	/organism="Mus musculus"		
	/strain="C57BL/6J"		
	/db_xref="taxon:10090"		
	/clone="IMAGE:871408"		
	/clone_id="Soares mouse NM_0010005"		
	/sex="male"		
	/tissue_type="heart"		
	/dev_stage="4 weeks"		
	/lab_host="Dbl0B"		
	/note="Vector: pTT73D-Pac (Pharmacia) with a modified		
	polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA		
	was purified with a Not I - oligo(dT) primer [5'-		
	TTTTTACCACTGAGGGGGCGCCGAGTTTTTTTTTTTTTTTTTTTTTT		
	3'-]; double-stranded cDNA was ligated to Eco RI adaptor		
	(Pharmacia), digested with Not I and cloned into the Not		
	I and Eco RI sites of the modified pTT73 vector. RNA		

BASE COUNT	Query Match Best Local Similarity	Score 20;	DB 9;	Length 405;
ORIGIN	Bonardo."	100.0%;	Pred. No. 16;	
QY	413 AATAGCTGCCATCCCTC 432	Matches : 20;	Conservative	Mismatches : 0;
Db	36 AATACTGCCCCATCCCC 55	Version : 4.1	Indels : 0;	Gaps : 0;
SOURCE				
ORGANISM	Mus musculus			
Locus	BB810434	408 bp	mRNA	linear EST 19-NOV-2001
DEFINITION	BB810434 RIKEN full-length enriched cDNA clone RCB-0558		LLC cDNA	Mus
REFERENCE	BB810434	MUSCULUS	CDNA clone G730005F13 3'	mRNA sequence.
AUTHORS	BB810434.1	GI:16983063		
VERSION		EST		
KEYWORDS		house mouse.		
TITLE				
JOURNAL				
COMMENT				
CONTACT	Toshihiko Hayashizaki			
	laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center(GSC), Yokohama Institute			
	The Institute of Physical and Chemical Research (RIKEN)			
	1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan			
TEL	81-45-503-9222			
FAX	81-45-503-9215			
EMAIL	genome-rss@gsc.riken.go.jp			
URL	http://genome.gsc.riken.go.jp/			
CARNICCI, P., SHIBATA, Y., HAYASHI, N., SUGAHARA, Y., SHIBATA, K., ITOH, M., KONNO, H., OKAZAKI, Y., MURAMATSU, M. and HAYASHIZAKI, Y.	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes.	Genome Res.	10 (10), 1617-1630 (2000)	
WAGI, K., FUJIWARA, S., INOUE, K., TOGAWA, Y., IZAWA, M., OHARA, E., WATABIKI, M., YONEDA, Y., ISHIKAWA, T., OZAWA, K., TANAKA, T., MATSUURA, S., KAWAI, J., OKAZAKI, Y., MURAMATSU, M., INOUE, Y., KIRA, A. and HAYASHIZAKI, Y.	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer.	Genome Res.	10 (11), 1757-1771 (2000)	
KONNO, H., FUKUNISHI, Y., SHIBATA, K., ITOH, M., CARNICCI, P., SUGAHARA, Y. and HAYASHIZAKI, Y.	Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library.	Genome Res.	11 (2), 281-289 (2001)	
	Please visit our web site (http://genome.gsc.riken.go.jp) for further details.			
FEATURES	mouse tissues	Location/Qualifiers		
SOURCE	1. 408	/organism="Mus musculus"		
	/ab_xref="taxon:10090"	/clone="G730005F13"		

/clone_lib="RIKEN full-length enriched, lung RCB-0558 LLC
 /DNA"
 /tissue_type="lung"
 /cell_line="RCB-0558 LLC"
 /note="pooled cell lines : (cell_line=CRL-1751 WEHI 164),
 (cell_line=CRL-2116 JC), (cell_line=RCB-035 WERT-3),
 (cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
 (cell_line=RCB-0559 K-1, Fl), (cell_line=RCB-1283 B-6
 melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
), (cell_type=Leydig cells, cell_line=CRL-2065 MITC-1),
 (cell_type=NULLipotent stem cell, cell_line=CRL-2070 NE),
 (tissue_type=bladder, cell_line=RCB-0544 MPM-2),
 (tissue_type=bone marrow, cell_type=stroma cell,
 cell_line=RCB-2028 (9114732141gb) AF129072.1), (tissue_type=kidney,
 cell_line=CCL-142 RAG), (tissue_type=submandibular gland,
 cell_line=CRL-1734 SCA-9 clone 15), (strain=BALB/c,
 cell_type=B cells, cell_line=RCB-1669 BC11 Clone 13 20-3B3
 BC3H)" strain="C3H", tissue_type="brain", cell_line=CRL-1443

BASE COUNT	ORIGIN	a	c	g	t
134		89	9	57	9
				128	t

RESULT 13
 A2327901/C
 LOCUS 1M051K01 R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
 DEFINITION clone UUGCLM051K05 R, DNA sequence.
 VERSION A2327901
 KEYWORD GSS
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE I (bases 1 to 413)
 AUTHORS Dunn,H., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Lonaopare,S., Mahmoud,M., Meenon,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingay,A., von Niederhäusern,A.,
 and Wright,D., Weiss,R.
 TITLE Mouse whole genome scaffolding with paired end reads from 10kb
 plasmid inserts
 COMMENT Unpublished (2000)
 JOURNAL Contact: Robert B. Weiss
 University of Utah Genome Center
 University of Utah, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
 Tel: 801 585 5606
 Fax: 801 585 7177
 Email: dunn@genetics.utah.edu
 Insert Length: 10000 Std Error: 0.00
 Plate: 0051 Row: K Column: 05
 Seq primer: CACACAGAAACRGCTATGACC
 Class: Plasmid ends
 High quality sequence stop: 413.
 FEATURES source
 I. 413
 /organism="Mus musculus"
 /strain="C5BL/6J"
 /db_xref="taxon:10090"
 /clone="UUGCLM051K05"
 /clone_1lb="Mouse 10kb plasmid UUGCLM library"
 /sex="Male"

/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
 /note="Vector: PMD42mv; purified genomic DNA from M.
 musculus C5BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (<http://www.jax.org/resources/documents/nihares/>). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA Polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (9114732141gb) AF129072.1, a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptorized mouse DNA was annealed to
 adaptored vector DNA, and transformed into
 chemically-competent *E. coli* XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

BASE COUNT	ORIGIN	a	c	g	t
91		91	a	75	9
				147	t

RESULT 14
 BF225508/C
 LOCUS BF225508
 DEFINITION UY41f07.x1 NCI_CGAP_Lu30 Mus musculus cDNA clone IMAGE:3662149 3',
 mRNA sequence.
 VERSION BF225508
 KEYWORDS EST
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE I (bases 1 to 418)
 AUTHORS Dunn,H., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
 Islam,H., Lonaopare,S., Mahmoud,M., Meenon,E., Pedersen,T., Reilly,
 M., Rose,M., Rose,R., Stokes,R., Tingay,A., von Niederhäusern,A.,
 and Wright,D., Weiss,R.
 TITLE Unpublished (1997)
 JOURNAL Other ESTs: uylf07.y1
 CONTACT: Robert Strasberg, Ph.D.
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 CGNA Library Preparation: Life Technologies, Inc.
 CGNA Library Arrayed by: The I.M.A.G.E. Consortium (LINE)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LINE at:
<http://line.lnl.gov/image/html/resources.shtml>
 MGI:1422917
 High quality sequence stop: 406.
 FEATURES source
 I. 418
 /organism="Mus musculus"
 /strain="CZCH II"
 /clone="IMAGE:3662149"
 /clone_1lb="NCI_CGAP_Lu30"
 /tissue_type="tumor, metastatic to mammary"
 /lab_host="BL10B"
 /note="Organ: lung; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: SalI; transgenic model WNT-1, expression driven by

MMTV-LTR enhancer; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigator providing samples: Gilbert Smith, NIHⁿ.
ORIGIN BASE COUNT 130 a 60 c 88 g 140 t
DEFINITION Query Match 4.2%; Score 20; DB 13; Length 420;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
KEYWORDS QY 413 AATACCTGCCATTCCCCC 432
DB 387 ATACCTGCCATTCCCCC 368

RESULT 15
BG975401
LOCUS BG975401 60284052F1 NCL_CGAP_Mam4 Mus musculus mRNA linear EST 12-JUN-2001
DEFINITION mRNA sequence.
ACCESSION BG975401
VERSION BG975401.1 GI:14363038
KEYWORDS EST
SOURCE house mouse.
ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 420)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgsbs-r@mail.nih.gov
Tissue Procurement: Lothar Hennighausen, Ph.D., Priscilla Furth
 Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov/plate:lim10975> row: h column: 13
 High quality sequence stop: 418.

FEATURES
source
 1..420
 /organism="Mus musculus"
 /strain="NMR1"
 /db_xref="Raxon:10000"
 /clone="IMAGE:4978620"
 /clone.lib="NCL_CGAP_Mam4"
 /tissue_type="tumor, gross tissue"
 /lab="DHL0B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Library constructed by Life Technologies. Primer: Oligo dT. Providing samples: Lothar Hennighausen/Priscilla Furth. NIH Reference for transgenic model: Li et al., Cell Growth and Differentiation 7, 3-11 (1995)."
BASE COUNT 149 a 87 c 58 g 125 t 1 others
ORIGIN
 Query Match 4.2%; Score 20; DB 13; Length 420;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 413 AATACCTGCCATTCCCCC 432
DB 22 ATACCTGCCATTCCCCC 41

Search completed: July 20, 2003, 07:09:58
 Job time : 1131 secs

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Copyright (c) 1993 - 2003 Compugen Ltd.
GenCore version 5.1.6

OM nucleic - nucleic search, using sw model

SUMMARIES

Run on: July 20, 2003, 03:28:35 ; Search time 1535 Seconds
(without alignments)
9005.754 Million cell updates/sec

Title: US-09-719-002-1
Perfect score: 475
Sequence: 1 gaatcttattttggacatgt.....acaacacacacaatcatgt 475

Scoring table: IDENTITY.NUC
Gapop 10.0 , gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Genbank:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_p1:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fn:*

17: em_hum:*

18: em_in:*

19: em_mnu:*

20: em_cm:*

21: em_or:*

22: em_cv:*

23: em_Pat:*

24: em_Ph:*

25: em_Pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_htq_hum:*

31: em_htq_inv:*

32: em_htq_other:*

33: em_htq_mus:*

34: em_htq_pln:*

35: em_htq_cod:*

36: em_htq_main:*

37: em_htq_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Result No.	Score	Query Match	Length	DB	ID	Description
1	451	94.9	473	6	AX008564	AX008564 Sequence
2	53.4	11.2	240020	2	ACI21587	ACI21587 Mus musculis
3	53.4	11.2	289723	2	ACI22935	ACI22935 Mus musculus
4	52	10.9	171361	2	ACI03286	ACI03286 Rattus norvegicus
5	50.2	10.6	219342	10	AL732443	AL732443 Mouse DNA
6	49.6	10.4	16228	6	AX252126	AX252126 Sequence
7	49.6	10.4	16228	6	AX448891	AX348891 Sequence
8	48.6	9.8	37973	6	AX37098	AX37098 Sequence
9	48.4	10.2	15649	6	AX277940	AX277940 Sequence
10	48.4	10.2	15649	6	AX233627	AX233627 Sequence
11	47.8	9.1	349980	6	AX44550	AX344550 Sequence
12	47.4	10.0	8044	6	AX251380	AX251380 Sequence
13	46.4	9.8	549	3	DDAC1	X16525 Dictyostelia
14	46.4	9.8	7522	6	AX45817	AX345817 Sequence
15	46.4	9.8	349980	6	AX34556	AX34556 Sequence
16	46.2	9.7	171146	9	AL161450	AL161450 Human DNA
17	46	9.7	5856	6	AX45471	AX345471 Sequence
18	46	9.7	6655	6	AX77841	AX277841 Sequence
19	46	9.7	6665	6	AX323516	AX323516 Sequence
20	46	9.7	6665	6	AX44985	AX344985 Sequence
21	46	9.7	101297	2	AC098899	AC098899 Rattus norvegicus
22	46	9.7	349980	5	AX34556	AX34556 Sequence
23	45	9.6	12592	6	AX251796	AX251796 Sequence
24	45.6	9.6	349980	6	AX44573	AX344573 Sequence
25	45.4	9.6	18997	6	AX37385	AX347385 Sequence
26	45.4	9.6	18997	6	AX449106	AX349106 Sequence
27	45.4	9.6	18997	6	AX45776	AX345776 Human DNA
28	45	9.5	105783	9	AL162276	AL162276 Human DNA
29	45	9.5	6172	6	AX344634	AX344634 Sequence
30	45	9.5	8883	6	AX251519	AX251519 Sequence
31	45	9.5	48068	9	AL189983	AL189983 Human DNA
32	45	9.5	216734	5	AB073376	AB073376 Oryzias latipes
33	44.8	9.4	5268	6	AX45779	AX345779 Sequence
34	44.8	9.4	17738	6	AX46440	AX346440 Sequence
35	44.8	9.4	321003	2	PEMALP3	AL035476 Plasmodium falciparum
36	44.6	9.4	7352	6	AX345272	AX345272 Sequence
37	44.6	9.4	139990	2	AC055599	AC055599 Rattus norvegicus
38	44.6	9.4	272410	2	ACI07492	ACI07492 Rattus norvegicus
39	44.6	9.4	349980	6	AX344555	AX344555 Sequence
40	44.4	9.3	1141	6	AX037744	AX037744 Sequence
41	44.4	9.3	349980	6	AX44557	AX344557 Sequence
42	44.2	9.3	5768	6	AX47064	AX347064 Sequence
43	44	9.3	4981	6	AX059550	AX059550 Sequence
44	44	9.3	63292	8	F14615	ATL47260 Arabidopsis thaliana
45	44	9.3	152619	2	AC101920	AC101920 Mus musculus

ALIGNMENTS

Pred. No. is the number of results predicted by chance to have a

JOURNAL	COMMENT
PATENT: WO 9966057-A 1 23-DEC-1999; DRAFTER JOHN (GB); KENTON PAUL (GB); BIOGEMMA UK LTD (GB); PAUL WYATT (GB)	
FEATURES	
"source	
1. 473 /organism="Asparagus officinalis" /db_xref="taxon:4686"	
BASE COUNT	162 a 112 c 67 g 132 t
ORIGIN	
Query Match Similarity 94.9% ; Score 451; DB 6; Length 473; Matches 473; Conservative 0; Mismatches 0; Indels 2; Gaps 2;	
QY	1 GAATTCTTATTCGACCTGACTCTGTGGCTGCCGAGTGCGTCCGAATTCTGT 60
DB	1 GAATTCTTATTCGACCTGACTCTGTGGCTGCCGAGTGCGTCCGAATTCTGT 60
QY	61 TGGCACACATCTGGCTCTCTGATGTTGAGCTGCATCCATAATTTCATGCA 120
DB	61 TGGCACACATCTGGCTCTCTGATGTTGAGCTGCATCCATAATTTCATGCA 120
QY	121 GAGAGACGACATGACTAAGTAATAGCTATACCCATAACTGATACAAGGAT 180
DB	121 GAGAGACGACATGACTAAGTAATAGCTATACCCATAACTGATACAAGGAT 180
QY	181 GAGACATCCACAGAAATACTCAATTAATAGCTGTGGTGTGAATGGAATGAA 240
DB	181 GAGACATCCACAGAAATACTCAATTAATAGCTGTGGTGTGAATGGAATGAA 240
QY	241 CTGATTAATGACTGAACTTCAAGAACTTCAACAGGCTAACATGAGCTAAT 300
DB	241 CTGATTAATGACTGAACTTCAAGAACTTCAACAGGCTAACATGAGCTAAT 300
QY	301 TCAACAGAAAGACTTCAAGAACTTCAACAGGCTAACATGAGCTAACATGAGCT 360
DB	301 TCAACAGAAAGACTTCAAGAACTTCAACAGGCTAACATGAGCTAACATGAGCT 360
QY	361 CAATTCGATATCCATGGACTGAGACCACTCAAATTATCCCTTAATTAACCTG 420
DB	361 CAATTCGATATCCATGGACTGAGACCACTCAAATTATCCCTTAATTAACCTG 419
QY	421 CCATTCCTCTCCAGACTCATCACAAACACACACACACATCATG 475
DB	420 CCATTCCTCTCCAGACTCATCACAAACACACACACACATCATG 473
RESULT 2	
AC121587	
LOCUS	AC121587 240020 bp DNA linear HTG 25-MAY-2002
DEFINITION	Mus musculus chromosome UNK clone RP23-27302, WORKING DRAFT
SEQUENCE	11 unordered pieces.
ACCESSION	AC121587
VERSION	2 GI:2106457
KWISWORDS	HTGS_PHASE1; HTGS_DRAFT.
SOURCE	
ORGANISM	Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	
REFERENCE	1 (bases 1 to 240020)
AUTHORS	McPherson, J.D. and Waterston, R.H.
JOURNAL	The sequence of <i>Mus musculus</i> clone Unpublished
REFERENCE	2 (bases 1 to 240020)
AUTHORS	McPherson, J.D. and Waterston, R.H.
JOURNAL	Direct Submission
REFERENCE	Submitted (20-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
AUTHORS	McPherson, J.D. and Waterston, R.H.
JOURNAL	Direct Submission
REFERENCE	3 (bases 1 to 240020)
AUTHORS	McPherson, J.D. and Waterston, R.H.
JOURNAL	Submitted (25-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
FEATURES	
"source	
1. 473 /organism="Mus musculus" /db_xref="taxon:1090" /chromosome="UNK" /clone="RP23-27302"	
misc_feature	
1. 1565 /note="assembly_name:Contig956"	
misc_feature	
1. 1565 /note="assembly_name:Contig957"	
misc_feature	
1. 1565 /note="assembly_name:Contig958"	
misc_feature	
1. 1565 /note="assembly_name:Contig99"	
misc_feature	
1. 1565 /note="assembly_name:Contig60"	
misc_feature	
1. 1565 /note="assembly_name:Contig61"	
misc_feature	
1. 1565 /note="assembly_name:Contig62"	

misc_feature 127310 . 155476
 /note="assembly_name:Contig63"
 misc_feature 155577 . 230360
 /note="assembly_name:Contig64"
 misc_feature 238461 . 239069
 /note="assembly_name:Contig933"
 BASE COUNT 71540 a /note="assembly_name:Contig925"
 ORIGIN a 49697 c 48772 g 68844 t 1167 others

Query Match 11.2%; Score 53.4; DB 2; Length 240020;
 Best Local Similarity 54.3%; Pred. No. 0.013; Mismatches 91; Indels 0; Gaps 0;
 Matches 108; Conservative 0; MisMatches 91; Indels 0; Gaps 0;

QY 167 AATGAAACAGGAGATGACCATCCAGAAAAATTCTAATTAGTCCTTGGCGTAGAAT 226
 Db 67358 AAAAGAGACATGATGATCATTGCAATTGTTCTTCTTGTGTCATT 67417
 QY 227 TGGAACTGATAGACTTAATTAATGACTTGTGCAATAATAAAAATGAAAGTCCTA 286
 Db 67418 TGTACAGAGATATTAAATATTAAATGATGCACTATTGAGATAATTGAA 67477
 QY 287 ACATGAAGACTAGTCTACATAGAGTAGCCACACTGACTTACATCAGAAG 346
 Db 67478 ATAAATAATATAATATAATATAATACAGATCACAGTACAGTAACAGAAG 67537
 QY 347 GCTTGTACCTTCACAAT 365
 Db 67538 GCGTAGTGTACTGAGAAT 67556

RESULT 3
 AC122935 AC122935 289723 bp DNA linear HTG 29-MAY-2002
 DEFINITION Mus musculus chromosome UNK clone RP23-20E4, WORKING DRAFT
 ACCESSION AC122935
 LOCUS Mus musculus chromosome UNK clone RP23-20E4
 VERSION 1.0
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE
 ORGANISM Mus musculus

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Buteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 289723)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE The sequence of *Mus musculus* clone
 JOURNAL Unpublished
 PUBLISHER 2 (bases 1 to 289723)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (28-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA
 REFERENCE 3 (bases 1 to 289723)
 AUTHORS McPherson, J.D. and Waterston, R.H.
 TITLE Direct Submission
 JOURNAL Submitted (29-MAY-2002) Genome Sequencing Center, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT ----- Genome Center -----
 Center: Washington University Genome Sequencing Center
 Center code: WUGSC
 Web site:<http://genome.wustl.edu/gsc/index.shtml>
 Contact: submissions@wuston.wustl.edu
 Center Project name: M_BA0020B04

----- Summary Statistics -----
 Sequencing vector: M13; 0%
 Sequencing vector: Plasmid; 100%
 Chemistry: Dye-primer ET; 0% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly Program: Phrap; version 0.990319

FEATURES Source
 misc_feature 1. 289723
 /organism="Mus musculus"
 /db_xref="taxon:0090"
 /chromosome="UNK"
 /clone="RP23-20E4"
 misc_feature 1. 3055
 /note="assembly_name:Contig9"
 misc_feature 3156 . 7772
 /note="assembly_name:Contig20"
 misc_feature 7873 . 11259
 /note="assembly_name:Contig21"
 misc_feature 11360 . 16743
 /note="assembly_name:Contig922"
 1684 . 2533
 misc_feature 2434 . 37441
 /note="assembly_name:Contig24"
 37542 . 73397
 /note="assembly_name:Contig25"
 73498 . 118271
 misc_feature 188372 . 217099
 /note="assembly_name:Contig26"
 misc_feature 217200 . 289723
 /note="assembly_name:Contig27"
 BASE COUNT 88915 a 57516 c 56681 g 85704 t 907 others
 ORIGIN

Query Match 11.2%; Score 53.4; DB 2; Length 289723;
 Best Local Similarity 54.3%; Pred. No. 0.013; Mismatches 91; Indels 0; Gaps 0;
 Matches 108; Conservative 0; MisMatches 91; Indels 0; Gaps 0;

QY 167 AATGAAACAGGAGATGACCATCCAGAAAAATTCTAATTAGTCCTTGGCGTAGAAT 226
 Db 65911 AAAAGAGACATGATGATCATTGCACTATTGTTCTTGTGTCATT 65970
 QY 227 TGGAACTGATAGACTTAATTAATGACTTGTGCAATAATTAGAGAAGTCCTA 286

Db	65971	TGTACAGAGGATTTAATTAAATTTCATGCCATTATTATGAA	66030
REFERENCE	Baylor Plaza, Houston, TX 77030, USA		
AUTHORS	3 (bases 1 to 171361)		
TITLE	Worley, K.C.		
JOURNAL	Direct Submission		
COMMENT	Submitted (13-JUL-2002) Human Genome Sequencing Center of Molecular and Human Genetics, Baylor College of Medicine, Baylor Plaza, Houston, TX 77030, USA On Jul 11, 2002 this sequence version replaced gi:17977		
Center: Baylor College of Medicine			
Center code: BCM			
Web site: http://www.hgsc.bcm.tmc.edu/			
Contact: hgsc-help@bcm.tmc.edu			
Project Information			
Center Project name: CH230-12H2			
Sequencing vector: Plasmid;			
Chemistry: Dye-terminator Big Dye: 100% of reads			
Assembly Program: Phrap; version 0.990329			
Consensus quality: 141160 bases at least Q30			
Consensus quality: 146679 bases at least Q20			
NOTE: Estimated insert size may differ from sequence length.			
(see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft.html)			
NOTE: This is a 'working draft' sequence. It currently			
consists of 59 contigs. The true order of the pieces			
is not known and their order in this sequence record			
arbitrary. Gaps between the contigs are represented			
runs of N, but the exact sizes of the gaps are unknown.			
This record will be updated with the finished sequence			
as soon as it is available and the accession number			
be preserved.			
1048: contig of 1048 bp in length			
1149: gap of unknown length			
2645: contig of 1498 bp in length			
2647: gap of unknown length			
2747: contig of 1085 bp in length			
3831: gap of unknown length			
3931: gap of unknown length			
3932: contig of 1375 bp in length			
5307: gap of unknown length			
5406: gap of 2246 bp in length			
5407: contig of 1481 bp in length			
7653: gap of unknown length			
7752: gap of 1185 bp in length			
7753: contig of 1185 bp in length			
8938: gap of unknown length			
9038: contig of 1402 bp in length			
10439: gap of 1599 bp in length			
10440: gap of unknown length			
10539: gap of unknown length			
10700: contig of 1227 bp in length			
10701: gap of unknown length			
12021: gap of unknown length			
12121: contig of 1444 bp in length			
13565: gap of unknown length			
13665: contig of 1599 bp in length			
13667: gap of unknown length			
15362: contig of 1599 bp in length			
15363: gap of unknown length			
16289: contig of 1227 bp in length			
16590: gap of unknown length			
16690: contig of 1217 bp in length			
17906: gap of unknown length			
17907: gap of unknown length			
18007: contig of 1739 bp in length			
19745: contig of 1739 bp in length			
19845: gap of unknown length			
21008: contig of 2063 bp in length			
19746: gap of unknown length			
19846: gap of unknown length			
21009: gap of unknown length			
22008: gap of unknown length			
22009: contig of 1155 bp in length			
23164: gap of unknown length			
23264: contig of 1189 bp in length			
24452: gap of unknown length			
24553: contig of 1154 bp in length			
27070: gap of unknown length			
25807: contig of 2116 bp in length			
27923: gap of unknown length			
28022: gap of unknown length			
28023: contig of 1631 bp in length			
29653: gap of unknown length			
29654: contig of 1416 bp in length			
31169: gap of unknown length			

Query Match 10.6%; Score 50.2; DB 10; length 219342;
 Best Local Similarity 54.6%; Pred. No. 0.073; Mismatches 83; Indels 0; Gaps 0;
 Matches 100; Conservative 0; Mismatches 83; Indels 0; Gaps 0;

DEFINITION Sequence 349 from Patent WO0202807.
 ACCESSION AX48891
 VERSION AX48891.1 GI:18614926
 KEYWORDS synthetic construct.

SOURCE ORGANISM synthetic construct.
 REFERENCE Olek,A., Piepenbrock,C. and Berlin,K
 AUTHORS Diagnoses of diseases associated with cell signalling
 TITLE Patent: WO 0202807-A 349 11-JAN-2002;
 JOURNAL /db.xref="taxon:34630"
 /note="chemically treated genomic DNA (Homo sapiens)"

QY 139 AAGGTTATGCTTAACTCCCTAAACTAATACAAACGGAGACATCCACAGAAA 198
 Db 69492 AAGGGATTAATTGTTAGTTCGCCACTCTAAATATGGAAACRAATGCCCACARACTA 69433

QY 199 ATTCGAATAGTCUTTGCGTGTGAAATGGAACGTGATACTACATTAATCAACT 258
 Db 69432 TTCTAAAGTAGTAATTCCTATTGCAATTGACAATAGAGATCAGTAGTTAGATGTC 69373

QY 259 TTGCAATTAAATAATAGAAAGCTAACATGAGAGACTAGTCTAACATGAGACTGT 318
 Db 69372 CCATATTAATTTAAATTAATCAACTATGATATGATCCATCTAACTGAGACTGA 69313

QY 319 CCA 321
 Db 69312 CAA 69310

RESULT 6
 AX252126/c
 LOCUS AX252126 16228 bp DNA linear PAT 05-OCT-2001
 DEFINITION Sequence 387 from Patent WO0168911.
 ACCESION AX252126
 VERSION 1 GI:15985481

KEYWORDS SOURCE ORGANISM
 synthetic construct.
 artificial sequences.

REFERENCE AUTHORS TITLE
 Olek,A., Piepenbrock,C. and Berlin,K
 Diagnosis of diseases associated with the cell cycle
 Patient: WO 0168911-A 387 20-SEP-2001;

FEATURES FEATURES
 Location/Qualifiers
 1. 16228

(ab_xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 4652 a 198 c 3601 g 7777 t
 ORIGIN

Query Match 10.4%; Score 49.6; DB 6; Length 16228;
 Best Local Similarity 48.3%; Pred. No. 0.15; Mismatches 139; Indels 0; Gaps 0;
 Matches 139; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

DEFINITION Sequence 2169 from Patent WO0200928.
 ACCESSION AX47098
 VERSION AX47098.1 GI:18494986

KEYWORDS SOURCE ORGANISM
 synthetic construct.
 artificial sequences.

REFERENCE AUTHORS TITLE
 Olek,A., Piepenbrock,C. and Berlin,K
 Diagnosis of diseases associated with the immune system
 Patient: WO 0200928-A 2169 03-JAN-2002;

FEATURES FEATURES
 Location/Qualifiers
 1. .37973
 /organism="synthetic construct"
 /db.xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 10166 a 620 c 8741 g 18446 t
 ORIGIN

Query Match 10.2%; Score 48.6; DB 6; Length 37973;
 Best Local Similarity 48.9%; Pred. No. 0.22; Mismatches 158; Indels 1; Gaps 1;
 Matches 158; Conservative 0; Mismatches 164; Indels 1; Gaps 1;

DEFINITION Sequence 349 from Patent WO0202807.
 ACCESSION AX48891
 VERSION AX48891.1 GI:18614926

KEYWORDS synthetic construct.

SOURCE ORGANISM synthetic construct.
 REFERENCE Olek,A., Piepenbrock,C. and Berlin,K
 AUTHORS Diagnoses of diseases associated with cell signalling
 TITLE Patent: WO 0202807-A 349 11-JAN-2002;
 JOURNAL /db.xref="taxon:34630"
 /note="chemically treated genomic DNA (Homo sapiens)"

QY 9001 TCACGAATCAAATAAAACCATCTAACACACATAATAAACCC 8954
 Db

RESULT 8
 AX347098/c
 LOCUS AX347098 37973 bp DNA linear PAT 01-FEB-2002
 DEFINITION Sequence 2169 from Patent WO0200928.
 ACCESION AX47098
 VERSION AX47098.1 GI:18494986

KEYWORDS SOURCE ORGANISM
 synthetic construct.
 artificial sequences.

REFERENCE AUTHORS TITLE
 Olek,A., Piepenbrock,C. and Berlin,K
 Diagnosis of diseases associated with the immune system
 Patient: WO 0200928-A 2169 03-JAN-2002;

FEATURES FEATURES
 Location/Qualifiers
 1. .37973
 /organism="synthetic construct"
 /db.xref="taxon:32630"
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 10166 a 620 c 8741 g 18446 t
 ORIGIN

Query Match 10.2%; Score 48.6; DB 6; Length 37973;
 Best Local Similarity 48.9%; Pred. No. 0.22; Mismatches 158; Indels 1; Gaps 1;
 Matches 158; Conservative 0; Mismatches 164; Indels 1; Gaps 1;

DEFINITION Sequence 349 from Patent WO0202807.
 ACCESSION AX48891
 VERSION AX48891.1 GI:18614926

KEYWORDS synthetic construct.

SOURCE ORGANISM synthetic construct.
 REFERENCE Olek,A., Piepenbrock,C. and Berlin,K
 AUTHORS Diagnoses of diseases associated with cell signalling
 TITLE Patent: WO 0202807-A 349 11-JAN-2002;
 JOURNAL /db.xref="taxon:34630"
 /note="chemically treated genomic DNA (Homo sapiens)"

QY 9001 TCACGAATCAAATAAAACCATCTAACACACATAATAAACCC 8954
 Db

RESULT 7
 AX348891/c
 LOCUS AX348891

Db 26628 AAATAACTTAAATTCATCAAAACAAAAAATTCATATAACACTA 26569
 QY 190 ACAGAAAAATCTTAATTAGTCCTTGCGTGTGAAATGGAACTGAACTTCA 249
 Db 26568 AAACAAAAATCACAAATTAACATAACATAACAAATCCGAAATCCTACATCA 26509
 QY 250 TTACAACTTGTCAAATAATAAACGAAAGTCTACATGAAGACTAGTTACATG 309
 Db 26508 CCGTAA-ACATCAACTAAATAATAATATACATACATACATACATACATAC 26450
 QY 310 AAGACTGAGTCCAGGAACTGGTACTTATCCCAAAGGCTTAGACTTCCACAATGAG 369
 Db 26449 TAACCAGTACAAATACACTGTTACATACATCCACCTTAAACCGAAAAGCTA 26390
 QY 370 ATTATCCATGGACTGAGGACACATCAATTCCTAAATACCTGCCATTC 429
 Db 26369 ATACAGAAATCACAAATAAGAACCATCTAAACATAAAACCCGTCCTCTA 26330
 QY 430 CTCTCCGACTGATCTACTCA 452
 Db 26329 AAATACAAATAATACCGCA 26307

RESULT 9
 AX277940/c
 LOCUS AX277940 DEFINITION Sequence 103 from Patent WO0177375. DNA
 ACCESSION AX277940 VERSION GI:15605012
 KEYWORDS SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
 TITLE Diagnoses of diseases associated with dna transcription
 JOURNAL Patent: WO 0192565-A 115 06-DEC-2001.
 FEATURES /organism="synthetic construct"
 source 1. .15649
 /note="chemically treated genomic DNA (Homo sapiens)"
 BASE COUNT 4082 a 261 c 3731 g 7575 t
 ORIGIN
 Query Match 10.2%; Score 48.4; DB 6; Length 15649;
 Best Local Similarity 46.7%; Pred. No. 0; 28; Mismatches 176; Indels 0; Gaps 0;
 Matches 154; Conservative 0; Mismatches 176; Indels 0; Gaps 0;
 Db 11798 CTAACACCAACCTACTCTCCAACTCCAAATCAATTTCATAAAACAA 11739
 QY 196 AAATCTTAATTAGCTTGCGTGAAGAATGAACTGAACTCATATACAA 255 -
 Db 11738 ACACACTTAACACTAACTATCTTAAATTAACATTAACATTAACATTAACATAATA 11679
 QY 256 CTTTGCATAAATATAAGAAGTCTAAATGAGACTGTCTAACAGAACT 315
 Db 11678 CATTACTTCTAAATTAACACATTAACACATCTCCAACTCCAAAT 11619
 QY 316 AGTCACGAACTGGAACCTATCCACAAAGGTAGCTTCACAAATGAGATT 375
 Db 11618 TTTCTTCCATTTTTAAACAAATTGCTTACCTCCAACTAAATACAA 11559
 QY 376 CCAGGACTGAGCACATCCAAATTCCTATAATACCTGCCATCCCTCAACCTCAACCTCAAAAT 11469
 Db 11558 ATAAGATATCTCACTCACTACACCTCGGCCATTAATTCCTACTT 1149
 QY 436 CAGACTCATCACTAAACACACACA 465
 Db 11498 CAACCTCTAAATACATAACATCAAA 11469

RESULT 11
 AX34450/c
 LOCUS AX344550 DEFINITION Sequence 1 from Patent WO0200932. DNA
 ACCESSION AX344550 VERSION GI:18492436
 KEYWORDS SOURCE synthetic construct.
 ORGANISM artificial sequences.
 REFERENCE AUTHORS Olek,A., Piepenbrock,C. and Berlin,K.
 TITLE Diagnoses of known genetic parameters within the mhc
 JOURNAL Patent: WO 0200932-A 1 03-JAN-2002;
 Epigenomics AG (DE)
 FEATURES Location/Qualifiers
 source 1. .34980
 /organism="synthetic construct"

/db_xref="taxon:32630"
/note="chemically treated genomic DNA (IGENOME)

GeCore version 5.1.6						
Copyright (c) 1993 - 2003 Compugen Ltd.						
OM nucleic - nucleic search, using sw model						
Run on: July 20, 2003, 03:02:54 ; Search time 191 Seconds						
{without alignments}						
5600.519 Million cell updates/sec						
Title: US-09-719-002-1	c	10	46	9.7	6665	22 AAS45299
Pefect score: 475	c	11	46	9.7	6665	24 ABL2083.
Sequence: 1 gaattctttatcgacactga.....acaacacacacaaatcatg 475	c	12	46	9.7	6665	24 ABK24130
Scoring table: IDENTTY_NUC	c	13	45.6	9.6	12592	24 AAS61101
gapop 10.0 , Gapext 1.0	c	14	45.4	9.6	18997	24 ABK31949
Searched: 2185239 seqs, 112599159 residues	c	15	45.4	9.6	18997	24 ABL22571
Total number of hits satisfying chosen parameters:	c	16	45	9.5	6172	24 ABB8042
Minimum DB seq length: 0	c	17	45	9.5	8883	22 AAS616761
Maximum DB seq length: 200000000	c	18	44.8	9.4	526	24 ABL32877
Post-processing: Minimum Match 0%	c	19	44.8	9.4	17738	24 ABL33538
Listing first 45 summaries	c	20	44.6	9.4	7352	24 ABL32370
Database : N_Geneseq_101022,*	c	21	44.2	9.3	5768	24 ABL34162
1: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1980.DAT:*	c	22	44	9.3	163319	21 AAF22306
2: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1981.DAT:*	c	23	43.8	9.2	9646	24 ABL33688
3: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1982.DAT:*	c	24	43.6	9.2	6162	24 ABL92315
4: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1983.DAT:*	c	25	43.6	9.2	14253	24 ABL33495
5: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1984.DAT:*	c	26	43.4	9.1	5557	24 ABL33549
6: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1985.DAT:*	c	27	43.4	9.1	8622	24 ABL34143
7: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1986.DAT:*	c	28	43.4	9.1	11944	24 ABL34187
8: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1987.DAT:*	c	29	43.2	9.1	1812	22 AAS5502
9: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1988.DAT:*	c	30	43.2	9.1	11812	22 AAS61742
10: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1989.DAT:*	c	31	43.2	9.1	11812	24 ABN00210
11: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1990.DAT:*	c	32	43.2	9.1	11812	24 ABK80432
12: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1991.DAT:*	c	33	42.8	9.0	6694	24 ABL33665
13: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1992.DAT:*	c	34	42.8	9.0	513445	22 AAL1373
14: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1993.DAT:*	c	35	42.6	9.0	3920	24 ABQ7030
15: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1994.DAT:*	c	36	42.6	9.0	6290	24 ABT33047
16: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1995.DAT:*	c	37	42.6	9.0	10717	24 ABN00210
17: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1996.DAT:*	c	38	42.5	9.0	10717	24 ABL33694
18: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1997.DAT:*	c	39	42.6	9.0	33053	24 ABQ70059
19: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1998.DAT:*	c	40	42.4	8.9	5958	24 ABL33665
20: /SIDS2/gcadata/geneseq/geneseq-emb1/NA1999.DAT:*	c	41	42.4	8.9	6866	24 ABT49320
21: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2000.DAT:*	c	42	42.4	8.9	6866	24 ABU32667
22: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2001A.DAT:*	c	43	42.4	8.9	15732	22 AB445389
23: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2002.DAT:*	c	44	42.4	8.9	15732	24 ABK28234
24: /SIDS2/gcadata/geneseq/geneseq-emb1/NA2002.DAT:*	c	45	42.4	8.9	83391	24 ABQ67094
RESULT 1						
ID: AAZ29510	XX	XX	XX	XX	XX	XX
AAZ29510 standard; DNA; 475 BP.						
AC: AAZ29510;	XX	XX	XX	XX	XX	XX
DT: 14-PAR-2000 (first entry)	XX	XX	XX	XX	XX	XX
A.officialis thaumatin-like PR-5 related gene (AOPRT-L) promoter.						
Inducible promoter; thaumatin-like PR-5 related gene; AOPRT-L; trait; KW non-phytotoxic inducing agent; Salicylic acid; SA; BTH; transformed plant; KW systemic activation; developmental; environmental; pathogen resistance; KW heterologous gene; disease control; sterility; fertility; fruit ripening; ds. disease control; sterility; fertility; KW Asparagus officinalis.						
SUMMARIES						
8						
Result No.	Score	Query Match Length	DB ID	Description	Location/Qualifiers	
c	1	475	100.0	A.officialis thaumatin-like PR-5 related gene (AOPRT-L) promoter.	misc_feature	100..111
c	-2	49.6	10.4	Chemically treated Human immune system	misc_feature	/*tag= a sequence="Homologous to Potato wound induced promoter
c	3	49.6	10.4	Chemically treated Human immune system	misc_feature	205..220 /*tag= b sequence="Homologous to Potato wound induced promoter sequence"
c	4	48.6	10.2	Chemically pretreated DNA transcription	misc_signal	226..341 /*tag= c label= SA_responsive_element
c	5	48.4	10.2	Tumour suppressor	misc_feature	/*tag= "Induced by non-phytotoxic agents like Salicylic acid or BTH"
c	6	47.4	10.0	Human immune system	misc_feature	251..264
c	7	46.4	9.8	Human immune system	misc_feature	
c	8	46	9.7	Human immune system	misc_feature	
c	9	5856	24	Human immune system	misc_feature	

FT /*tag= d
 FT /note= "Homologous to Carrot PR-3 and PR-4 promoter
 FT sequence"
 FT - repeat_region
 FT 281..318
 FT /*tag= e
 FT /rpt-type= TANDEM
 FT 281..298
 FT /*tag= f
 FT /note= "18 bp repeat"
 FT 409..416
 FT /*tag= g
 XX
 PN WO9966057-A2.
 XX
 PD 23-DEC-1999.
 XX
 XK 21-JUN-1999; 99WO-GB01949.
 PR 19-JUN-1998; 98GB-0013345.
 XX
 PA (BIOG-) BIOGEMMA UK LTD.
 XX
 PI Draper J, Kenton P, Paul W;
 XX
 DR WPI; 2000-106107/09.
 XX
 Novel promoters used to control the expression of heterologous genes in
 PT transformed plants -
 XX
 PS Sequence 475 BP; 163 A; 113 C; 67 G; 132 T; 0 other;
 XX
 Query Match 100.0%; Score 475; DB 21; Length 475;
 Best Local Similarity 100.0%; Pred. No. 1.6e-113;
 Matches 475; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GATTCTTATGGACCTGACTCTCTGTTGCTGCCAGGTGCTGTC
 1 GAATTCTTATGGACCTGACTCTCTGTTGCTGCCAGGTGCTGTC
 61 TCGCACACATACTGGCTCTGCTGATGACAGCTTCCATATA
 61 TGGCACACATACTGGCTCTGCTGATGACAGCTTCCATATA
 120 121 GAGAGAACGACATGACTAACTTGTAAATTAGCTTAA
 121 GAGAGAACGACATGACTAACTTGTAAATTAGCTTAA
 180 181 GACACATCCACGAAATACTTAATTAGCTTGGCTGAAATGG
 181 GACACATCCACGAAATACTTAATTAGCTTGGCTGAAATGG
 240 241 CTACATATGACAACTTGTCAAATTAAATAAGAAGTCT
 241 CTACATATGACAACTTGTCAAATTAAATAAGAAGTCT
 300 301 TCTACACGAGACTGCTCCAGCAACTGACTCTATT
 301 TCTACACGAGACTGCTCCAGCAACTGACTCTATT
 360 301 TCTACACGAGACTGCTCCAGCAACTGACTCTATT
 360
 QY 361 CAATCGAGATATCCATGGACTGACACCATCCAATTATC
 361 CAATCGAGATATCCATGGACTGACACCATCCAATTATC
 420 421 CCCATTCCTCCCTCAGACTCACTCAACACACAC
 475 421 CCCATTCCTCCCTCAGACTCACTCAACACACAC
 475
 FT TATA_signal
 FT /*tag= g
 RESULT 2/
 ID ABL70459/C
 ID AB170459 standard; DNA; 16228 BP.
 XX
 AC ABL70459;
 XX
 DT 01-JUL-2002 (first entry)
 XX
 DE Chemically treated cell signalling DNA sequence#175.
 XX
 KW Cell signalling; cytosine methylation; cell signalling disease;
 KW cancer; tumour; cytostatic; ds.
 XX
 OS Unidentified.
 XX
 PN WO200202807-A2.
 XX
 PD 10-JAN-2002.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPT; 2002-154758/20.
 XX
 PS Claim 1; SEQ ID NO 349; 24pp+sequence listing; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence of at least
 CC 18 bases of a segment of chemically pretreated DNA of genes associated
 with cell signalling. The activity of the modified sequences of the
 CC invention may be described as cytostatic. The object of the invention is
 CC to provide the chemically modified DNA of genes associated with cell
 CC signalling, as well as oligonucleotides and/or RNA-oligomers for
 CC detecting cytosine methylations, as well as a method which is
 CC particularly suitable for the diagnosis and/or therapy of genetic and
 CC epigenetic parameters of genes associated with cell signalling. The
 CC chemically modified DNA provided by the invention is useful for diagnosis
 CC and therapy of diseases such as solid tumours and cancer. The sequences
 CC given in records ABL70111-AB170526 represent chemically pre-treated
 CC genomic DNA's of genes associated with cell signalling.
 Note: The sequence data for this patient is not represented in the printed
 CC specification, but is based on sequence information supplied by the
 European Patent Office.
 XX
 Sequence 16228 BP; 4652 A; 198 C; 3601 G; 7777 T; 0 other;
 SQ
 Query Match 10.4%; Score 49.6; DB 24; Length 16228;
 Best Local Similarity 48.3%; Pred. No. 0.0064;
 Matches 139; Conservative 0; Mismatches 149; Indels 0; Gaps 0;
 QY 131 CATGACTTAACTGATAGCTTAACTCTACACAGATGAC
 131 CATGACTTAACTGATAGCTTAACTCTACACAGATGAC
 190 9241 CTTTACATTTACACAACTTAAATTAAA
 9241 CTTTACATTTACACAACTTAAATTAAA
 9182 191 CAGAAAAATTCTTAATTAGCTTGTGCTGAAATGAA
 191 CAGAAAAATTCTTAATTAGCTTGTGCTGAAATGAA
 250

PT for diagnosis and treatment of diseases associated with abnormal

PT cytosine methylation -

XX X

PS - Claim 1; SEQ ID NO 2169; 32pp + sequence Listing; German.

XX C

GCC The present invention provides a number of human immune system associated

CC genes which are modified by the methylation of cytosines. The sequences

CC can be used in the diagnosis and treatment of immune system disorders,

CC - including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid

CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,

CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel

CC diseases. The present sequence is a gene of the invention.

XX Sequence 37973 BP; 10166 A; 620 C; 8741 G; 18446 T; 0 other;

XX Query Match Best Local Similarity 10.2%; Score 48.6; DB 24; Length 37973;

XX Matches 158; Conservative 0; Mismatches 164; Indels 1; Gaps 1;

QY 130 ACATGACTAAGTATTAGCTTAACTCCATAACAGAGATCACCC 189

Db 26628 AAATRACTAAATAATCATCAAAACAAACAAAAAAATTCATACACACTA 26559

QY 190 ACAGAAAATTCGAATAGCTTGCTGAACTGATACCTACATTA 249

Db 26568 AAAACAAAAATAACAAATAATTAACATACAAATCGAAAATCTACTATCA 26509

QY 250 TTACAACTTGGACAAATAATTAAGAAGTCTACATGAGACTGTCATACG 309

Db 26508 CGGTAACATCAAACTAAATAAAATTATTAACATACATACACAA 26450

QY 310 AAGACTGTTCCAGGACTGTGACTTATCCAAAGGCTTAGACTTCCACAAATCGAG 369

Db 26449 TAACCGAAATACATAACCGTACTATATCCACACCTTAAACACG 26390

QY 370 ATATCCCATGGACTGAGACACCAATAATCCTAAACACATATAAACCCCCTCTCTCA 26330

Db 26369 ATACGAAATCAAATACTGAACCATCTAACACATATAAACCCCTCTCTCA 26390

QY 430 CTCTCCGAGCTACATCACTCA 452

Db 26329 AATATACAAATTAATTAACGCA 26307

RESULT 5

AS45396/C AS45396 standard; DNA; 15649 BP.

AC XX AS45396;

XX 18-DEC-2001 (first entry)

DE Chemically pretreated genomic DNA associated with cell cycle #51.

XX Cell cycle; human; Cpg dinucleotide; cytosine methylation; HIV; aging;

XX human immunodeficiency virus; neurodegenerative disorder; solid tumour;

XX graft-versus-host disease; glomerular disease; Levy body disease; cancer;

XX arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarrhythmic;

XX immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;

XX PCR primer.

XX Homo sapiens.

XX WO200168911-A2.

XX 20-SEP-2001.

XX 15-MAR-2001; 2000DE-1013847.

PR 15-MAR-2000; 2000DE-1019058.

PR 06-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1048826.

XX XX

PA (EPIC-) EPIGENOMICS AG.

XX XX

PI Olek A, Piepenbrock C, Berlin K;

DR WPI; 2001-602751/68.

PT Designing primers and probes for analysing diseases associated with

PT cytosine methylation state e.g. arthritis, cancer, aging,

PT arteriosclerosis comprising fragments of chemically modified genes

XX associated with cell cycle.

XX PS Claim 1; SEQ ID NO 101; 28pp; English.

XX Sequences AS45296-AS45520 represent chemically pretreated genomic DNA

CC molecules associated with the cell cycle and specific PCR primers of the

CC invention. The sequences are useful for detecting the methylation state

CC of all CpG dinucleotides in a sequence and therefore for analysing

CC associated diseases. By analysing cytosine methylations in the pretreated

CC DNA, generic and/or epigenetic parameters for the diagnosis and therapy

CC of existing diseases or the predisposition to specific diseases can be

CC ascertained. The parameters may be compared to another set of genetic

CC and/or epigenetic parameters, the differences serving as basis for

CC diagnosis and/or prognosis events which are disadvantageous to patients.

CC The sequences of the invention are useful for the diagnosis and therapy

CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,

CC aging, glomerular disease, lewy body disease, arthritis,

CC arteriosclerosis, solid tumours and cancers.

XX Sequence 15649 BP; 4082 A; 261 C; 3731 G; 7575 T; 0 other;

XX Query Match Best Local Similarity 10.2%; Score 48.4; DB 22; Length 15649;

XX Matches 154; Conservative 0; Mismatches 176; Indels 0; Gaps 0;

QY 136 CTAAGTAATTAGCTTAATCCCTAAACTCATACAAACGAGATCACAGAA 195

Db 11798 CTAACAAACCAAACCTACTCTAACCTCCAAACATTAATTITTCATAAAAAAACA 11739

QY 196 AAATCTTAATTAGCTTCTGGTGTAGAAATTGAACTGATACTACATCATATACAA 255

Db 11738 ACAAACTTAAACACTAAACTATCTTAAATAAATTAACAACTAA 11679

QY 256 CTTTGCAAAATAATTAAGAAGTCTAACTGAACTGAGCTTCACAAATCGAGATTC 315

Db 11678 CATTACTTCTTAATTTACAGATTAACCACTTCAACCTTCAACCTTCAACAT 11619

QY 316 AGTCACGAACTCTGAACTTATCCACAAAGGCTTAGCTTCACAAATCGAGATTC 375

Db 11618 TTTCCTTCATTTTTTTAACTAAATTGCTCTATTCGCCTTCCAACTAAATACA 11559

QY 376 CCAGGAGCTGATGGACACATCCAAATPATCCCTATAATACTGGCOATCCTCCCTC 435

Db 11558 ATAAGGATATCTCACTACTACATACCTACCCGCCATTAATTCACATTCCTT 11499

QY 436 CAGCTCTACTACAAACACACACAC 465

Db 11498 CAACCTCTAAATACATACACATA 11469

RESULT 6

ABK2841/C ABK2841 standard; DNA; 15649 BP.

XX XX

AC ABK2841;

XX DT 23-APR-2002 (first entry)

DE DNA transcription associated genomic DNA #58.

XX XX

KW DNA transcription associated gene; peptide nucleic acid; PNA-oligonomer;

KW PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KW single nucleotide polymorphism; adenosine deaminase deficiency; cancer;
 KW viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KW immunological disorder; Werner syndrome; developmental disorder;
 KW neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KW myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KW angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
 KW polyglutamine disorder; solid tumour.
 OS Unidentified.

XX WO200192565-A2.

PN 06-APR-2001; 2001WO-EP03973.

PD 06-DEC-2001.

XX (EPIC-) EPIGENOMICS AG.

XX 06-APR-2000; 2000DE-1019058.

PR 07-APR-2000; 2000DE-1019173.

PR 30-JUN-2000; 2000DE-102529.

PR 01-SEP-2000; 2000DE-1043826.

XX PA

XX PI Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-090046/12.

XX New nucleic acids or oligomers, useful for diagnosing or treating
 PT diseases associated with DNA transcription, e.g. immunological
 PT disorders, Werner syndrome, psoriasis, myocardial infarction, solid
 PT tumours or cancer

XX PS Claim 1; SEQ ID NO 115; 32pp; English.

CC The invention relates to a nucleic acid, which comprises a segment of the
 CC chemically pretreated DNA of genes associated with DNA transcription from
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
 CC to the chemically pretreated DNA of genes associated with DNA
 CC transcription. The set of oligomer probes are useful for detecting the
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
 CC diagnosing or treating diseases associated with DNA transcription
 CC (particularly with the methylation status), e.g. adenosine deaminase
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,
 CC haematological disorders, immunological disorders, Werner syndrome,
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
 CC neurological disorders, neurodegenerative disorders, Waardenburg
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
 CC infarction, hypertension, angiogenesis, erythropoiesis, congenital heart
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
 CC or cancer. Sequences ABK28127-ABK2872 represent DNA transcription
 CC associated genomic DNA molecules of the invention.

CC Note: The sequence data did not form part of the printed
 CC specification but was obtained in electronic format directly from the
 European Patent Office.

XX Sequence 15649 BP; 4082 A; 261 C; 3731 G; 7575 T; 0 other;

XX SO Query Match 10.2%; Score 48.4; DB 24; Length 15649;

XX Best Local Similarity 46.7%; Pred. No. 0_013; Mismatches 154; Conservativity 0; Matches 154; Indels 0; Gaps 0;

QY 136 CTTAAGTATTAACTCTTAATCCCTAAACTCATACAAACGAGATCACATCCACAGAA 195

Db 11798 CTAAACACCAACACTACTCTCAACTCCCAAATTCATTTTCAAAACACCA 11739

QY 196 AAAATTCTAACTAGTCCTTGCGGTAGAAATGGAACTGATACTCATTAATCACA 255

Db 11738 ACAAACTTAAACACTAACATCTATCTTAATTAATTAATAAACTAAATA 11679

OY 256 CTTTGCAAATAAATATAAGAAAGTCTAACATGAAGACTAGTCATACATGAGACT 315

Db 11678 CATTAACTTCTTAACTATATAACATTTAACCATCTCCACCTCAAACCTCAAAT 11619

OY 316 AGTCAGGAATGCGACTTTCACAGAGCTTACATCCAAATGGAACT 375

Db 11618 TTTCCTTCCATTTCATTTTAACAAATTGCGCTCATTCGCCAAACTAAATACA 11559

OY 376 CCAAGGACTGAGACCATCAAATTATCCCTATAATACCTGGCCATCCGCCCT 435

Db 11558 ATACCCATATCACTACATACACCTCGGCCTTAATTCAACATTCTCAGT 11499

OY 436 CAGACATCATCTACTGAAACAAACACACA 465

Db 11498 CAACCTCTAAATACTAAACTAAACTACAAATA 11469

DT RESULT 7

XX RAS4626/C

XX ID AAS4626 standard; DNA; 8044 BP.

XX PR AAS4626;

XX AC

XX XX

XX DT 18-DEC-2001 (first entry)

XX DE Tumour suppressor gene derived chemically modified sequence #348.

XX PR Human; tumour suppressor gene; oncogene; antitumour; cytostatic;

XX KW cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;

XX KW cytosine methylation; ds.

XX OS Homo sapiens.

XX PR

XX PN WO200168912-A2.

XX PD 20-SEP-2001.

XX PR 15-MAR-2001; 2001WO-EP02955.

XX PT Fragments of chemically modified genes associated with tumour suppressor
 PT genes and oncogenes, useful in designing primers and probes for
 PT analysing diseases associated with cytosine methylation state e.g.
 PT cancer

XX PS Claim 1; SEQ ID NO 348; 27pp; English.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPI; 2001-602752/6B.

XX PT

CC The invention relates to a nucleic acid comprising a sequence of 18
 CC bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
 CC bisulphite, of genes associated with tumour suppression and
 CC oncogenes having a sequence taken from 536 (actually 533 since
 CC numbers 408, 458 and 500 are missing from the sequence listing) sequences
 CC (S_s) and sequences complementary to (S_s). The nucleic acid may be a
 CC peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may
 CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the

CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes. Sequences with even numbered Seq ID numbers are the
 CC complementary sequence of the corresponding odd numbered sequence (e.g.
 CC ID 2 and IDL, ID 536 and ID 535, except for those whose partner sequence
 CC is missing).

Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC [ftp.wipo.int/pub/published_pct_sequences](http://wipo.int/pub/published_pct_sequences).

SQ

Sequence 8044 BP; 2017 A; 211 C; 1923 G; 3893 T; 0 other;

Query Match 10.0%; Score 47.4; DB 22; Length 8044;
 Best Local Similarity 47.9%; Pred. No. 0.02; Indels 0; Mismatches 168;
 Matches 181; Conservative 0; Gaps 1; Del 1; Insert 2; PT
 124 AGAGGACATGACTAAGTAATTGCTTAACTCAATCACACAGATGAC 183
 Db 1413 AAATCATACATAAAATATAAAATATAAAATACATCACAAATAATTAAC 1354

QY 184 ACATCCACAGAAATACTTATTAGCTTGCGGTAGAAATGGAACTGAATCCTA 243
 Db 1353 ATTAAATAATAATTTATAATCTTATTTATTTATTTATTTATTTATTTA 1294

QY 244 CAATTAATACAACATTGCAATAATAATATAAGAAGTCAATGAGACTGTT 303
 Db 1293 CAATATACAACTATCTACCTAAATAACTCTTAACTTAACTTAACTTAA 1234

QY 304 AACATGAGACTAGTCAGCAGACTCTAACCAAGGCTTACATTCCACAA 363
 Db 1233 ACTTATACCGATAAGTAACTCTACGCCATAATCACACCTTAAACCGAAAT 1174

QY 364 ATGGAGATATGCCAAGGACGGA - TGGACACCCTCAATTATCCCTAAATACCTGC 421
 Db 1173 AAACAAATTACCTAAATAAAATTAACACTCTTACCAACATAAAAAACCTAT 1114

QY 422 CCTTCCCTCTCCAGACTCATCTRACTCHAACACACACACACACATC 472
 Db 1113 CTCTACTAAATACTACAAATTAACCAATATAACACACCTATAATC 1063

RESULT 8

ABL32915/C
 ID ABL32915 standard; DNA; 7522 BP.

XX ABL32915;
 AC ABL32915;

DT 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 888.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antianteriosclerotic; anisotropic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

KW Homo sapiens.

XX Homo sapiens.

PN WO200200928-A2.

XX 02-JUL-2001; 2001WO-EP07537.

PR 30-JUN-2000; 2000DE-1032529.

PR 01-SEP-2000; 2000DE-1043826.

PA (EPIC-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K;

XX DR WPT; 2002-130909/17.

XX PT Nucleic acid comprising fragment of chemically modified gene, useful
 PT for diagnosis and treatment of diseases associated with abnormal
 PT cytosine methylation

XX PS Claim 1; SEQ ID NO 888; 32pp + sequence listing; German.

The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration; arteriosclerosis; anaemia; cancer; acute myeloid
 CC leukaemia; Alzheimer's disease; AIDS; epilepsy; neurofibromatosis;
 CC rheumatoid arthritis; psoriasis and inflammatory/ulcerative bowel
 CC diseases. The present sequence is a gene of the invention.

XX Sequence 7522 BP; 2074 A; 79 C; 1578 G; 3791 T; 0 other;

XX Query Match 9.8%; Score 46.4; DB 24; Length 7522;
 Best Local Similarity 53.3%; Pred. No. 0.035; Indels 0; Gaps 0;
 Matches 98; Conservative 0; Mismatches 0; Del 0; Insert 0; PT

QY 103 TAATTTCCAGTCATGAGAGCATGACTTAAGTAATTGCTTAATCCCTAA 162
 Db 3895 TAACCTTCTCTCTCATACACTATACCTTCAAAACAAACAAATCTCTTAAATC 3837

QY 163 ACTCATAACAGGAGCACATCCACAGAAAATCTTAAATTGCTTGGTAG 222
 Db 3836 TTTCACAAATATCTATAACATCAAATAATATCTTAAACAAATCTTAAATC 3777

QY 223 AAATGGAACTGTAACTCTACCTACATTAACTACATTGCAAAATAATAAGAAT 282
 Db 3776 ATATPACACAAAAAAACTTCTATACTTAACTCTAACTTAAACACCGT 3717

QY 283 TCTA 286
 Db 3716 TCTA 3713

RESULT 9

ABL3256/C
 ID ABL3256 standard; DNA; 5856 BP.

XX AC ABL3256;

XX DT 26-MAR-2002 (first entry)

XX Human immune system associated gene SEQ ID NO: 542.

KW Human; immune system disease; cytosine methylation; antiasthmatic;
 KW antianteriosclerotic; anisotropic; cytostatic; nootropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antiheumatic; antiarthritic; antidiabetic; antipsoriatic;
 KW antinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
 KW gene; ds.

KW Homo sapiens.

XX Homo sapiens.

PN WO200200928-A2.

XX 02-JUL-2001; 2001WO-EP07537.

XX 30-JUN-2000; 2000DE-1032529.

XX 01-SEP-2000; 2000DE-1043826.

XX (EPIG-) EPIGENOMICS AG.
 XX PD 20-SEP-2001.
 XX PT 15-MAR-2001; 2001WO-EP02945.
 XX PR 15-MAR-2000; 2000DE-1013847.
 XX PR 06-APR-2000; 2000DE-1019058.
 XX PR 07-APR-2000; 2000DE-1019173.
 XX PR 30-JUN-2000; 2000DE-1032529.
 XX PR 01-SEP-2000; 2000DE-1043826.

PS (EPIG-) EPIGENOMICS AG.
 XX PI Olek A, Piepenbrock C, Berlin K;
 XX DR WPI; 2002-130909/17.

PT Nucleic acid comprising fragment of chemically modified gene, useful for diagnosis and treatment of diseases associated with abnormal cytosine methylation

XX Claim 1; SEQ ID NO 542; 32pp + Sequence Listing; German.

CC genes which are modified by the methylation of cytosines. The sequences can be used in the diagnosis and treatment of immune system disorders, including eye diseases such as retinopathy, neovascular glaucoma and macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis, rheumatoid arthritis, psoriasis and inflammatory/vulcerative bowel diseases. The present sequence is a gene of the invention.

XX Sequence 5856 BP; 1482 A; 177 C; 1495 G; 2702 T; 0 other;

Query Match 9.7%; Score 46; DB 24; Length 5856;
 Best Local Similarity 46.8%; Pred. No. 0.042; Matches 145; Conservative 0; Mismatches 165; Indels 0; Gaps 0;

CC 160 AAACTCAATACAACAGGATGACATCCAGAAGAAAATCTTATTAGTCATGCTTGCGTG 219
 CC 5523 AAATAAAACAAAAACCTTCAATCTAACATCA 5484

CC 220 TGAATTTGGAATGACTTACATTAATCTACTTGTGCAAATAATAAGAA 279
 CC 5483 TATAATCCTTCCGACACACAAACCCATCTTCTACAAATAATAACAA 5424

CC 280 AGTTCTTACAGAAGACTAGTCTAACATGAGACTPAGTCACGACTGCTACCTATC 339
 CC 5423 AAAAATAAAATAACATCACCCAAACCATATATCCAACTCTCTTTTTTTTA 5364

CC 340 CACAAGGCTTAGACTTCCAAATGGAGATTCCGACTCTCTTTTTTTTA 5304

CC 5363 AACAAATCTCTTATACCAAACTAAATAGATACGGCATCTGCTACTACA 5304

CC 400 ATTATCCCTATAATACCTGCCCCATCCCTCCAGACTCATCTAAACAA 459
 CC 5303 AACCTCTCCGAAATTAGCCATCTCTACCTCAACTCTGAAATAACTAACTA 5244

CC 460 CACACACCA 469
 CC 5243 CAATACCCA 5234

XX RESULT 10
 XX ASN4599/C
 XX ASN4599 standard; DNA; 6665 BP.

AC ASN45299;
 XX AC ASN45299;
 XX DT 18-DEC-2001 { first entry}

DE - Chemically pretreated complementary DNA associated with cell cycle #2.

XX Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arteriosclerosis; anti-HIV; neuroprotective; antiarthritis; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.

XX Homo sapiens.

XX Human; immune system disease; cytosine methylation; antiasthmatic;
 PN WO200168911-A2.

XX RESULT 11
 AB132083/C
 ID AB132083 standard; DNA; 6665 BP.
 XX AC AB132083;
 XX DT 26-MAR-2002 { first entry}

DE Human immune system disease; cytosine methylation; antiasthmatic;
 KW Human; immune system disease; cytosine methylation; antiasthmatic;

CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 12592 BP; 3783 A; 195 C; 2736 G; 5878 T; 0 other;
 SQ Best Local Similarity 49.6%; Score 45.6; DB 24; Length 12592;
 Matches 144; Conservative 0; Mismatches 144; Indels 2; Gaps 1;

QY 208 AGCTTSCGTAGAAATTGAACTACATTAACCTTGCAATA 267
 CC | | | | | | | | | | | | | | | | | | | | | | | | | |
 CC ATACAAGAATTAATTAACCTAATTAATTAACCTATCATCTCAACTTAAACTA 2989
 DB 3048 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 QY 268 AAATATAAGAAGTCACATGAA 293
 CC ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 CC 2988 AAATAAAAAAACTAACCTAAACCTAA 2963

RESULT 13
 AA61101/C
 ID AA61101 standard; DNA; 12592 BP.
 XX
 AC AA61101;
 XX
 DT 29-JAN-2002 (first entry)
 XX
 DE Human gene regulation-associated gene oligonucleotide #56.
 XX
 KW Human; Gene regulation-associated gene; severe combined immunodeficiency;
 KW cardiac damage; inflammatory response; Haemophilia; Werner syndrome;
 KW asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;
 KW renal disease; Preeclampsia; cardiac allograft vascular disease;
 KW colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;
 KW immunosimulant; cardiant; antiinflammatory; coagulant; antiasthmatic;
 KW nephrotropic; gynecological; anti-tumour; immunosuppressive; cyrostatic;
 XX OS Homo sapiens.
 PN WO20017375-A2.
 XX
 PD 18-OCT-2001.
 XX
 PF 06-APR-2001; 2001WO-EPO3968.
 XX
 PR 06-APR-2000; 2000DE-1019058.
 PR 07-APR-2000; 2000DE-1019173.
 PR 30-JUN-2000; 2000DE-1032229.
 PR 01-SEP-2000; 2000DE-1043826.

^aPA (EPIG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 DR WBI; 2002-017470/02.
 XX
 PN New nucleic acid sequences from chemically modified genes associated
 PT with gene regulation, useful for analysing cytosine methylation associated
 PT with diseases -
 PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
 XX
 PS Claim 1; SEQ ID No 57; 26pp; English.
 XX
 CC The invention relates to 224 nucleic acid sequences comprising at least
 CC 18 bases of a chemically pretreated gene associated with gene regulation
 CC selected from 43 known genes (or complementary sequences). The
 CC chemical pretreatment converts cytosine bases unmethylated at the
 CC 5'-position to uracil or another base with hybridisation behaviour
 CC dissimilar to cytosine, to enable analysis of cytosine methylations.
 CC The DNA sequences, oligomers (or sets/arrays) and method are
 CC useful in the diagnosis of diseases (or predisposition to diseases)
 CC associated with gene regulation and in therapy of such diseases, by
 CC enabling analysis of the cytosine methylation patterns of such genes,
 CC kits are provided. They are especially useful in diagnosis
 CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
 CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
 CC asthma, HDR syndrome, Sathre-Chotzen syndrome, renal disease,
 CC preeclampsia, graft versus-host disease. The present sequence is a
 CC sequence included in the sequence data for this specification and is
 CC associated with the human gene regulation-associated genes.
 Note: The sequence data for this patient did not form part
 CC of the printed specification, but was obtained in electronic
 CC

RESULT 14
 ABK33949/C
 ID ABK33949 standard; DNA; 18997 BP.
 XX
 AC ABK33949;
 XX
 DT 18-JUN-2002 (first entry)
 XX
 DE Human DNA for staging of Astrocytomas, complement, #16.
 XX
 KW Human; ds; astrocytoma; cytostatic; staging; cysteine methylation; CPG;
 KW bisulphite; brain tissue; MALDI; ESI; electron spray mass spectrometry;
 KW matrix assisted laser desorption/ionization mass spectrometry.
 XX OS Homo sapiens.
 XX
 PN WO200202808-A2.
 XX
 PD 10-JAN-2002.
 XX
 PF 02-JUL-2001; 2001WO-EPO7538.
 XX
 PR 30-JUN-2000; 2000DE-1032529.
 PR 01-SEP-2000; 2000DE-1043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX Olek A, Piepenbrock C, Berlin K;
 DR WBI; 2002-171649/22.
 XX
 PN Novel chemically modified genomic DNA sequences, useful in the
 PT characterisation, classification, differentiation, grading, staging,
 PT treatment and/or diagnosis of astrocytomas or predisposition to
 PT astrocytomas.
 XX
 PS Claim 1; SEQ ID No 32; 37pp; English.
 XX
 CC The invention relates to a nucleic acid comprising a sequence (I) of at
 CC least 18 bases in length of a segment of chemically pre-treated genomic
 CC DNA which has any one of the sequences of (ABK33915-ABK34032) or its
 CC complement. Also included are an oligonucleotide or peptide nucleic
 acid (or set thereof) of at least 9 nucleotides which hybridises to (I),

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: July 20, 2003, 03:30:14 ; Search time 44 Seconds
(without alignments)

Title: US-09-719-002-1
Perfect score: 475

Sequence: 1 gaattttatcgacgtga.....acacacacaatcatg 475

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters:

882724

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_NB:*

1: /rgn2_6/pctdata/1/ina5A_COMBO.seq: *
2: /ccn2_6/pctdata/1/ina5B_COMBO.seq: *
3: /ccn2_6/pctdata/1/ina6A_COMBO.seq: *
4: /rgn2_6/pctdata/1/ina6B_COMBO.seq: *
5: /ccn2_6/pctdata/1/ina6C_COMBO.seq: *
6: /ccn2_6/pctdata/1/ina6backfileseq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match Length DB ID Description

1 38 8.0 1263 4 US-09-302-681-4 Sequence 4, Appli

2 36.8 7.7 6243 2 US-09-302-681-4 Sequence 1, Appli

3 36.4 7.7 4033 2 US-09-302-681-4 Sequence 5, Appli

4 35 7.4 3815 4 US-09-302-681-4 Sequence 196, App

5 34.4 7.2 457 4 US-09-302-681-4 Sequence 306, APP

6 34 7.1 2495 1 US-09-302-681-4 Sequence 307, APP

7 33.6 7.1 4673 1 US-07-638-431-1 Sequence 1, Appli

8 33.6 7.1 4673 5 PCT-US2000018-1 Sequence 1, Appli

9 33.6 7.1 9412 2 US-08-055-188-1 Sequence 1, Appli

10 33.6 7.1 168575 4 US-09-426-290-1 Sequence 1, Appli

11 33.2 7.0 1253 2 US-08-511-620-1 Sequence 1, Appli

12 33.2 7.0 8700 2 US-08-922-625-16 Sequence 1, Appli

13 33.2 7.0 8700 2 US-08-466-961A-16 Sequence 1, Appli

14 33.2 7.0 8700 2 US-08-645-133B-18 Sequence 12, Appli

15 33 6.9 357 2 US-08-666-405-12 Sequence 12, Appli

16 33 6.9 575 2 US-08-569-166-29 Sequence 29, Appli

17 33 6.9 882 2 US-08-666-405-27 Sequence 27, Appli

18 32.8 6.9 246240 2 US-08-724-394A-20 Sequence 20, Appli

19 32.8 6.9 246240 2 US-08-724-394A-21 Sequence 21, Appli

20 32.8 6.9 246240 2 US-08-724-394A-22 Sequence 22, Appli

21 32.6 6.9 1001 4 US-09-611-630-599 Sequence 598, Appli

22 32.6 6.9 1001 4 US-09-611-630-599 Sequence 599, Appli

23 32.6 6.9 12793 4 US-09-048-830-124 Sequence 124, Appli

24 32.6 6.9 15062 4 US-09-044-830-89 Sequence 89, Appli

25 32.6 6.9 19124 2 US-08-487-826B-13 Sequence 13, Appli

26 32.4 6.8 359 2 US-08-587-080-65 Sequence 65, Appli

27 32.4 6.8 5609 4 US-09-313-677-14 Sequence 14, Appli

28 32.4 6.8 5609 4 US-09-313-677-15 Sequence 15, Appli
29 32.4 6.8 7026 4 US-09-313-677-20 Sequence 20, Appli
30 32.4 6.8 7344 4 US-09-313-677-16 Sequence 16, Appli
31 32.4 6.8 98844 4 US-09-791-231-10 Sequence 10, Appli
32 32 6.7 600 4 US-09-385-982-482 Sequence 482, Appli
33 32 6.7 773 4 US-09-385-982-541 Sequence 541, Appli
34 32 6.7 2639 4 US-09-374-335-1 Sequence 1, Appli
35 31.8 6.7 246240 2 US-08-724-394A-20 Sequence 20, Appli
36 31.8 6.7 246240 2 US-08-724-394A-21 Sequence 21, Appli
37 31.8 6.7 246240 2 US-08-724-394A-22 Sequence 22, Appli
38 31.6 6.7 658 4 US-08-998-16-5995 Sequence 595, Appli
39 31.6 6.7 731 1 US-08-451-05A-2 Sequence 2, Appli
40 31.6 6.7 3558 4 US-09-134-001C-2439 Sequence 2439, Appli
41 31.6 6.7 13737 4 US-09-538-414-10 Sequence 10, Appli
42 31.4 6.6 568 4 US-08-642-274D-23 Sequence 23, Appli
43 31.4 6.6 568 4 US-09-952-01C-23 Sequence 23, Appli
44 31.4 6.6 2539 3 US-08-749-522-3 Sequence 3, Appli
45 31.4 6.6 5099 4 US-09-610-040-5 Sequence 5, Appli

STREET: 1 South Pinckney Street
CITY: CHARLESTON
STATE: SC ZIP: 29403
COUNTRY: USA
TELEPHONE: (803) 747-1000
FAX: (803) 747-1001
EMAIL: charleston@compugen.com
ADDRESS: Quarles & Brady
CITY: CHARLESTON
STATE: SC ZIP: 29403
COUNTRY: USA
TELEPHONE: (803) 747-1000
FAX: (803) 747-1001
EMAIL: charleston@quarles.com

RESULT 1
US-09-302-681-4
; Sequence 4, Application US/09302681
; Patent No. 6441149
; GENERAL INFORMATION:
; APPLICANT: Herrnstadt, Corrina
; APPLICANT: Ghosh, Sountra S.
; APPLICANT: Cleaver, William
; APPLICANT: Fahy, Eoin F.
; APPLICANT: Davis, Robert E.
; TITLE OF INVENTION: DIAGNOSTIC METHOD BASED ON FILE REFERENCE: 56008841011
; CURRENT APPLICATION NUMBER: US/09/302,681
; CURRENT FILING DATE: 1999-04-30
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: HOMO sapien
; US-09-302-681-4

Query Match 8.0%; Score 38; DB 4; Length 1263;
Best Local Similarity 57.6%; Pred. No. 0.1%;
Matches 68; Conservative 0; Mismatches 50; Indels 0; Gaps 0;
QY 102 ATATTAATTCCATGAGAGACATGACTAAGTAACTGCCTAA 161
Db 803 ATGATATGACATCATATGTGCAATATGAATAATTAAAGGTAATCCCAAAT 862
QY 162 ACTATATACACAGATACACATCCAGAACAAATCTAACTAGCTTGCGTG 219
Db 863 AAATAAAAACAGAACATTAAANGAGAGAGAGAGATGAGATGGCGTTGGGTG 920

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US08/9336,165A
 FILING DATE: 24-SEP-1997
 CLASSIFICATION: 536
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 60/027,032
 FILING DATE: 24-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Gimmi, Edward R.
 REGISTRATION NUMBER: 38,891
 REFERENCE/DOCKET NUMBER: P50549
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 610-270-4478
 TELEFAX: 610-270-5090
 TELEX:
 INFORMATION FOR SEQ ID NO: 196:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3815 base pairs
 TYPE: nucleic acid
 STRANEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: Genomic DNA

US-08-936-165A-196

RESULT 5
 US-09-643-597-306

Query Match Best Local Similarity 59.6%; Score 35; DB 4; Length 3815;
 Matches 59; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 AAAATTCTAATTAGCTTGGGTGAGAAATGGAACTGATACTACATTAATTCAA 255
 Db 2467 ATATATTCCTAGAACATCATCATTATTAAAGATAAGCTAGCGATATTAC 253
 QY 254 AACTTTGCAATAAAATAAAAGAGTCTAACATGA 292
 Db 2407 CAATTTCACAATAAAATAAAATTACGCA 2369

RESULT 6
 5428147-1/c

Patent No. 5428147
 APPLICANT: BARKER, RICHARD F.; KEMP, JOHN D.
 TITLE OF INVENTION: OCTOPINE T-DNA PROMOTERS
 NUMBER OF SEQUENCES: 17
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US08/91,538
 FILING DATE: 13-JUL-1993
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 869,216
 FILING DATE: 13-APR-1992
 APPLICATION NUMBER: 869,216
 FILING DATE: 13-APR-1992
 APPLICATION NUMBER: 440,432
 FILING DATE: 21-NOV-1989
 APPLICATION NUMBER: 553,786
 FILING DATE: 19-NOV-1983
 APPLICATION NUMBER: 741,034
 FILING DATE: 06-AUG-1991
 APPLICATION NUMBER: 144,775
 FILING DATE: 20-JAN-1988
 APPLICATION NUMBER: 485,614
 FILING DATE: 15-APR-1983
 APPLICATION NUMBER: 713,624
 FILING DATE: 10-JUN-1991
 APPLICATION NUMBER: 260,574
 FILING DATE: 21-OCT-1988
 APPLICATION NUMBER: 848,733
 FILING DATE: 01-APR-1986
 APPLICATION NUMBER: 535,354
 FILING DATE: 26-SEP-1983
 ; SEQ ID NO:1;
 ; LENGTH: 24595
 ; 5428147-1

Query Match Best Local Similarity 51.0%; Score 33.8; DB 6; Length 24595;
 Matches 80; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 200 TCTCTAATTAGCTTGGGTGAGAAATGGAACTGATACTACATTAATTCAA 259
 Db 23880 TCCCTTCAGGTGGCTATGAGTGAATGGTCTCCAAAATTAACACGA 23821
 QY 260 TCCAATAAAATAAAAGAGTCTACATGAGTGAATGGTCTACATGAGACTGC 319
 Db 23820 ACACACACAAACAAATACAGAAATTAGCAATAAATGAATTACAACCGCATATCCTG 23761
 QY 320 CACGAACTCGTACCTTATGCCAACAGCTTGA 356
 Db 23760 CCATCACTCACACAAATCACACAAATPAATGGTTT 23724

RESULT 7
 US-07-638-431-1/c

Sequence 1, Application US/07638431
 ; Patent No. 5198535
 GENERAL INFORMATION:
 APPLICANT: Hoffman, Stephen L.
 APPLICANT: Charoenvit, Yupon
 APPLICANT: Hestrom, Richard
 APPLICANT: Khushni, Srisin
 APPLICANT: Rogers IV, William O.

Query Match Best Local Similarity 54.9%; Score 34.4; DB 4; Length 457;
 Matches 89; Conservative 0; Mismatches 71; Indels 2; Gaps 1;

QY 138 AAAGTAAATTAGCTTACCTTAAACACTAACACAAACAGGATGAGCATCCA-CAGAA 195
 Db 9 AGGTAAATAGTTTAAATTAAGACAGATGCAACGAAAGAGAATGAGCTGC 68

Db 4247 GAAATTAATTGGAAAATGA 4228

RESULT 9

US-09-955-138-1/C

; Sequence 1, Application US/08955138A

; GENERAL INFORMATION:

; APPLICANT: Lefebvre, Daniel D.

; TITLE OF INVENTION: PLANT PHOSPHATASES

; FILE REFERENCE: PPL07-01

; CURRENT APPLICATION NUMBER: US/08/955-138A

; CURRENT FILING DATE: 1997-10-21

; NUMBER OF SEQ ID NOS: 119

; SOFTWARE: FastSEQ for Windows Version 3.0

; SEQ ID NO 1 LENGTH: 9412

; TYPE: DNA

; ORGANISM: SOLANUM TUBEROSUM

; FEATURE:

; NAME/KEY: VARIANT

; LOCATION: (1)...(17)

; OTHER INFORMATION: Xaa = Any Amino Acid

US-08-955-138-1

Query Match 7.1%; Score 33.6; DB 2; Length 9412;

Best Local Similarity 51.5%; Pred. No. 4;

Matches 101; Conservative 0; Mismatches 94; Indels 1; Gaps 1;

; SEQ ID NO 1 LENGTH: 168575

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (21181)...(21403)

; NAME/KEY: CDS

; LOCATION: (95252)...(95430)

; NAME/KEY: CDS

; LOCATION: (101753)...(101996)

; NAME/KEY: CDS

; LOCATION: (110324)...(110439)

; SEQ ID NO 1 LENGTH: 168575

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (110439)...(110524)

; NAME/KEY: CDS

; LOCATION: (110524)...(110639)

; SEQ ID NO 1 LENGTH: 1253 base pairs

; LENGTH: 1253 base pairs

; NAME/KEY: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

; LOCATION: (124058)...(124278)

; NAME/KEY: CDS

; LOCATION: (124278)

; NAME/KEY: CDS

; LOCATION: (127009)...(127130)

; NAME/KEY: CDS

; LOCATION: (128910)...(129139)

; SEQ ID NO 1 LENGTH: 168575;

; TYPE: DNA

; ORGANISM: Homo Sapiens

; FEATURE: CDS

TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 STRAIN: Nicotiana tabacum NN
 DEVELOPMENTAL STAGE: TMV-induced
 IMMEDIATE SOURCE:
 CLOSE: Cluster-A
 NAME/KEY: CDS
 LOCATION: 14..1126
 OTHER INFORMATION: /partial
 US-08-591-029-1

Query Match Score 7.0%; Score 33.2; DB 2; Length 1253;
 Best Local Similarity 53.0%; Pred. No. 2..6; Indels 0; Gaps 0;
 Matches 71; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

Qy 54 TTTCGTGCGCACACATACTGGTCTTGTTGATTGAGCAGTCCATAATTATTC 113
 Db 1232 TTITGGAGCACACANTTATGGTTTATATTGTATAGTICACATATCTGTTGC 1173
 Qy 114 ATGTCATGAGAAAGCACTAACTAACTAACTAACTCAATATCTGTTGC 1173
 Db 1172 TTAATTAAAGTAGTATACATGACACATTAGTAACTCCATCACTCATCTTGAG 1113
 Qy 174 ACGAGATGACAT 187
 Db 1112 ATGACACTCCCCAT 1099

RESULT 12
 US-08-392-625-16/c
 Sequence 16, Application US/08392625
 Patent No. 5837485

GENERAL INFORMATION:
 APPLICANT: Entian, Karl-Dieter
 APPLICANT: Gtz, Friederich
 APPLICANT: Schnell, No. 5837485bert
 APPLICANT: Augustin, Johannes
 APPLICANT: Engelke, Germar
 APPLICANT: Rosenstein, Ralf
 APPLICANT: Kalletta, Corina
 APPLICANT: Klein, Corina
 APPLICANT: Kupke, Thomas
 APPLICANT: Jung, G nthter
 APPLICANT: Kellner, Roland
 TITLE OF INVENTION: Biosynthetic Process For The Preparation
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue
 CITY: Washington
 STATE: D.C.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DCS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/392,625
 FILING DATE:
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/876,791
 FILING DATE: 30-APR-1992

ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 0652..0980002
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2000
 TELEX/FAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8700 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 US-08-392-625-16

Query Match Score 7.0%; Score 33.2; DB 2; Length 8700;
 Best Local Similarity 47.6%; Pred. No. 5;
 Matches 98; Conservative 0; Mismatches 108; Indels 0; Gaps 0;

Qy 128 GCATGACTAAAGTAATTAGCTTAACTAACAAACGAGATGACATT 187
 Db 8306 GAAGTCCTAAATGAAATAATGATAATGCTAAACACTAATGCCAAACTT 8247
 Qy 188 CCACAGAAAAAAATTCTAATTAGCTTGTGCTGTTGAAATTGAAACTGTTACATT 247
 Db 8246 TTAAGAATCCATAAACAAAAAAATATCGAGTTAACCTAACTACATA 8187
 Qy 248 AATACACTTTGCAAATAAAATATAAGAAAGCTAACATGAGACTAGTCTAAC 307
 Db 8186 CAGCACAAATTAAAATGAAATCAAAATTAAGGTTAACATGAGACTAGTCTAAC 307
 Qy 308 TGAGACTAGTCAGAACTCGTAC 333
 Db 8126 TCAGTTGTAAATCCACCTGTCC 8101

RESULT 13
 US-08-446-961A-16/c
 Sequence 16, Application US/08466961A
 Patent No. 5843709

GENERAL INFORMATION:
 APPLICANT: Entian, Karl-Dieter
 APPLICANT: Gtz, Friederich
 APPLICANT: Schnell, No. 5843709bert
 APPLICANT: Augustin, Johannes
 APPLICANT: Engelke, Germar
 APPLICANT: Rosenstein, Ralf
 APPLICANT: Kalletta, Corina
 APPLICANT: Klein, Corina
 APPLICANT: Kupke, Thomas
 APPLICANT: Wieland, Bernd
 APPLICANT: Jung, G nthter
 APPLICANT: Kellner, Roland
 TITLE OF INVENTION: Biosynthetic Process for the Preparation of
 CHEMICAL COMPOUNDS
 NUMBER OF SEQUENCES: 42
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Fox
 STREET: 1100 New York Avenue
 CITY: Washington
 STATE: D.C.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DCS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/466,961A
 FILING DATE: 06-JUN-1995
 CLASSIFICATION: 435
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: 435
 FILING DATE:
 PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/392,625
 FILING DATE: 22-FEB-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/676,791
 FILING DATE: 30-APR-1992
 PRIORITY APPLICATION DATA:
 APPLICATION NUMBER: US 07/784,234
 FILING DATE: 31-OCT-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,895
 REFERENCE/DOCKET NUMBER: 0852-0980
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2640
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 16:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 8700 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 FS-08-4666-961A-16

Query Match 7.0% Score 33
 Best Local Similarity 47.68; Pred. No
 Matches 99; Conservative 0; Mismatches 98

128	GCACATGACTAAAGTAAATTAGCTTAATC		
8306	GAATGTCATAAAATGAAATAAGTTA		
188	CCACAGAAAAAATCTTAATTAGCTTGTG		
8246	TTRAGATTCATTAACAAAAAAATAAT		
248	AATTACAATTTGCAATTAAATAATAAA		
8186	CAGCACAAATTAAAACATGCCAATTCAA		
308	TGAAGACTTAATCCCGAACTCGTAC	3	
8126	TCAAGTTGTAATCCCTACGTGTC	8	

RESULT 14
 FS-08-645-193B-18/C
 Sequence 18, Application US/08645193B
 General Information:
 Patent No. 5962253
 APPLICANT: Kupke, Thomas
 APPLICANT: Gotz, Friedrich
 APPLICANT: Kempfer, Christoph
 APPLICANT: Jang, Gunther
 TITLE OF INVENTION: Oxidative Decarboxylation of Invention: Catalyzed by Flavoproteins
 NUMBER OF SEQUENCES: 70
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Sterne, Kessler, Goldstein & Hegerman LLP
 STREET: 1100 New York Avenue, Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: U.S.A.
 ZIP: 20005
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version 1.0
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/645,193B
 FILING DATE: 13-MAY-1996
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 19 683
REFERENCE/DOCKET NUMBER: 102.164
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 661-8000
TELEFAX: (212) 661-8002
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 357 base pairs
TYPE: nucleic acid
STRANDEDNESS: Single
TOPOLOGY: Linear
HYPOTHETICAL: No
ORIGINAL SOURCE:
ORGANISM: Clostridium perfringens
FEATURE:
NAME/KEY: PMRP126
US-08-666-405-12

	Query Match	Score	DB 2;	Length
Best Local Matches	6.9%	33;		357;
Local Similarity	54.5%	Pred. No.	2;	
Matches	66; Conservative	0;	Mismatches	55; Indels
		0;	Gaps	0;

Qy 141 G T A A T T A G C T T A A T C C C U P A A A C T C A T A C A A C G A T G A C A T C A C A G A A A A A T 200
D b 43 G T A T A T G G T C A A A T C C T G A T T C G C T G A C A A A T A G A T T A T A G A T T A T A G G G A A 102

Qy 201 T C T A A T T A G I C T T I C G T G A G A A T T G S A A C T G A A T C C T A C A T T A T T A C A C T T 260
D b 103 T C T R I C T A A T A T A G G A T T A A A C T T A A A C T T A C C T G C A T T A G A T T A G G T A T T 162

Qy 261 G 261
D b 163 G 163

Search completed: July 20, 2003, 05:20:04
Job time : 46 secs

Gencore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

M nucleic - nucleic search, using sw model

run on: July 20, 2003, 05:19:14 ; Search time 171 Seconds
(without alignments)

scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

title: US-09-719-002-1
perfect score: 475
sequence:
1 gaatttttattggacatcg.....acaacacaaatccatcg 475

searched: 1439767 seqs, 1031500376 residues

total number of hits satisfying chosen parameters: 2879534

minimum DB seq length: 0
maximum DB seq length: 2000000000

post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

database : Published Applications NA:
 1: /cgn2_6/podata/1/pubnpa/us08_pubcomb.seq;
 2: /cgn2_6/podata/1/pubnpa/pcl_new_pub.seq;*
 3: /cgn2_6/podata/1/pubnpa/us05_new_pub.seq;*
 4: /cgn2_6/podata/1/pubnpa/us05_pubcomb.seq;*
 5: /cgn2_6/podata/1/pubnpa/us01_new_pub.seq;*
 6: /cgn2_6/podata/1/pubnpa/pctns_pubcomb.seq;*
 7: /cgn2_6/podata/1/pubnpa/us08_new_pub.seq;*
 8: /cgn2_6/podata/1/pubnpa/us08_pubcomb.seq;*
 9: /cgn2_6/podata/1/pubnpa/us09_new_pub.seq;*
 10: /cgn2_6/podata/1/pubnpa/us09_new_pub.seq;*
 11: /cgn2_6/podata/1/pubnpa/us09_new_pub.seq2;*
 12: /cgn2_6/podata/1/pubnpa/us09_pubcomb.seq;*
 13: /cgn2_6/podata/1/pubnpa/us08_pubcomb.seq;*
 14: /cgn2_6/podata/1/pubnpa/us10_new_pub.seq;*
 15: /cgn2_6/podata/1/pubnpa/us61_new_pub.seq;*
 16: /cgn2_6/podata/1/pubnpa/us60_new_pub.seq;*
 17: /cgn2_6/podata/1/pubnpa/us60_pubcomb.seq;*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

ALIGNMENTS

RESULT 1
 ; Publication No. US20030082609A1.
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; ATTORNEY: PIPPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; ATTORNEY: BURG, Michael
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
 ; FILE REFERENCE: 5013.1003
 ; CURRENT APPLICATION NUMBER: US10/239,676
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03968
 ; DE 10019058.8
 ; DE 10019173.8
 ; DE 1003259.7
 ; DE 10043226.1
 ; PRIORITY DATE: 2001-04-06
 ; 2000-04-06
 ; 2000-06-30
 ; NUMBER OF SEQ ID NOS: 228
 ; SEQ ID NO 103
 ; LENGTH: 15649
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-239-676-103

No.	Result	Query	Match	Length	DB	ID	Description
C 1	48.4	10.2	15649	15	US-10-239-676-103		Sequence 103, App
C 2	4.6	9.7	6665	15	US-10-239-676-4		Sequence 4, Appl
C 3	45.4	9.6	18997	15	US-10-172-086-18		Sequence 18, Appl
C 4	43.2	9.1	11812	15	US-10-139-076-20		Sequence 210, Appl
C 5	42.8	9.0	51309	12	US-09-754-053-A4		Sequence 4, Appl
C 6	42.4	8.9	15732	15	US-10-139-676-96		Sequence 96, Appl
C 7	41.8	8.8	4604	15	US-10-239-576-20		Sequence 200, Appl
C 8	40.8	8.6	18997	15	US-10-172-086-17		Sequence 17, Appl
C 9	40.2	8.5	6158	15	US-10-139-676-24		Sequence 24, Appl
C 10	39.8	8.4	7195	15	US-10-139-676-30		Sequence 30, Appl
C 11	39.4	8.4	414	11	US-09-754-053-A4		Sequence 6528, Appl
C 12	39.8	8.4	6050	15	US-10-172-086-50		Sequence 50, Appl
C 13	39.4	8.3	64061	11	US-09-190-988-1		Sequence 1, Appl
C 14	39	8.2	6167	15	US-10-139-676-41		Sequence 41, Appl
C 15	38.8	8.2	864	11	US-09-238-042-A-290		Sequence 4290, Appl

SUMMARIES

8

No.	Result	Score	Match	Length	DB	ID	
C 1	48.4	10.2	15649	15	US-10-239-676-103		Sequence 103, App
C 2	4.6	9.7	6665	15	US-10-239-676-4		Sequence 4, Appl
C 3	45.4	9.6	18997	15	US-10-172-086-18		Sequence 18, Appl
C 4	43.2	9.1	11812	15	US-10-139-076-20		Sequence 210, Appl
C 5	42.8	9.0	51309	12	US-09-754-053-A4		Sequence 4, Appl
C 6	42.4	8.9	15732	15	US-10-139-676-96		Sequence 96, Appl
C 7	41.8	8.8	4604	15	US-10-239-576-20		Sequence 200, Appl
C 8	40.8	8.6	18997	15	US-10-172-086-17		Sequence 17, Appl
C 9	40.2	8.5	6158	15	US-10-139-676-24		Sequence 24, Appl
C 10	39.8	8.4	7195	15	US-10-139-676-30		Sequence 30, Appl
C 11	39.4	8.4	414	11	US-09-754-053-A4		Sequence 6528, Appl
C 12	39.8	8.4	6050	15	US-10-172-086-50		Sequence 50, Appl
C 13	39.4	8.3	64061	11	US-09-190-988-1		Sequence 1, Appl
C 14	39	8.2	6167	15	US-10-139-676-41		Sequence 41, Appl
C 15	38.8	8.2	864	11	US-09-238-042-A-290		Sequence 4290, Appl

Qy Db

Sequence 123, App
Sequence 187, App
Sequence 43, App
Sequence 35, App
Sequence 204, App
Sequence 72, App
Sequence 3, Appl
Sequence 212736, App
Sequence 14, App
Sequence 95, App
Sequence 109062, App
Sequence 109063, App
Sequence 709, App
Sequence 135, App
Sequence 23, App
Sequence 6, App
Sequence 3400, App
Sequence 214763, App
Sequence 214761, App
Sequence 211, App
Sequence 52, App
Sequence 13, App
Sequence 121, App
Sequence 118, App
Sequence 9, App
Sequence 68, App
Sequence 191, App
Sequence 31, App

Oy 196 AAAATTCTAATTAGTGTTCGGTGTAGAAATTGGAACCTGAAATCCATTAATTAGAA 255
 Db 11738 ACAAAACTTAAACACTAAACTATCTTAAATTAAACCTAAATAATAATAATAATA 11679
 Qy 256 CTTTGCAATTAAATAAGAAAGTCTAACATGAAACTAGTCTAACATGAAAGACT 315
 Db 11678 CATTAACTTCAATTATACATTAACATTAAACCTCAACCTAAATAATAATA 11619
 Qy 316 AGTCCACGAACTGACTCATTCFACAAAGGCTAGACTTTCCACAAATCGATATT 375
 Db 11618 TTTCCTCCATTTTTAAACAAAATTGCTTATCCCAACTAAATAACA 11559
 Qy 376 CCATGGACTGTGGCACCATCCAAATTATCCCTATAATACCTGCCATTCCCTCTC 435
 Db 11558 ATAACGATATCTCAACTACATACACCTCCGCCTCTAAATTCAACAACTCTT 11499
 Qy 436 CAGACTOCTAACCTAACAAACAGACACA 465
 Db 11498 CAACCCCTAAATAACAAACTACAAATA 11469

RESULT 3
 US-10-172-086-18/C
 ; Sequence 18, Application US/10172086
 ; Publication No. US20030113750A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Epigenomics AG
 ; TITLE OF INVENTION: Method and nucleic acids for the differentiation
 ; of prostate tumors
 ; FILE REFERENCE:
 ; CURRENT APPLICATION NUMBER: US/10/172,086
 ; CURRENT FILING DATE: 2002-06-13
 ; NUMBER OF SEQ ID NOS: 116
 ; SEQ ID NO 18
 ; LENGTH: 18997
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-172-086-18

Query Match 9 6%; Score 45.4; DB 15; Length 18997;
 Best Local Similarity 49.0%; Pred. No. 0.31; DB 15;
 Matches 121; Conservative 0; Mismatches 0; Gaps 0;
 Indels 0;
 Qy 99 CCAATAATTTCATGCTAGGAGAAAGCACATGACTAAAGTAAATTAGCTTAACCCCC 158
 Db 11885 CAAACCTATTAAACCCGTTAACTAAATAATAACAACTAAACTACCTTTC 11826
 Qy 159 TAAACCTAAATCAACGAGATGACATCCACAGAAAAATTCTTAACTCTTGCGT 218
 Db 11825 TCCTCTCAATAAAACTAAACTATCATTAAACCTAAACAT 11766
 Qy 219 GTAGAAATTGGAAACTGAAACTACATTAACCTTGCACAAATAATAAAAGA 278
 Db 11765 TAAATAAAACCAATAAAATAAAATTCTCAATCTTAAATAAAATAAA 11706
 Qy 279 AACTCTAACATGAAAGACTAGTCTAACATGAAAGACTAGTCCACGAACTGTACCTTATI 338
 Db 11705 AACTATAAAATAAACACTAAACTATCATTAAACCTAAATCTTAAATTAC 11646
 Qy 339 CCACAAA 345
 Db 11645 ACAAAA 11639

RESULT 4
 US-10-239-676-4/C
 ; Sequence 4, Application US/10239676
 ; Publication No. US20030082609A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
 ; FILE REFERENCE: 5013.1003
 ; CURRENT APPLICATION NUMBER: US/10/239,676
 ; PRIORITY APPLICATION NUMBER: PCT/EP01/03968
 ; CURRENT FILING DATE: 2002-09-24
 ; PRIORITY FILING DATE: 2001-04-06
 ; PRIORITY APPLICATION NUMBER: DE1019058.8
 ; DE 10019173.8
 ; DE 10032529.7
 ; DE 10043826.1
 ; PRIORITY FILING DATE: 2001-04-06
 ; LENGTH: 6665
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; US-10-239-676-4

Query Match 9 7%; Score 46; DB 15; Length 6665;
 Best Local Similarity 51.5%; Pred. No. 0.13; DB 15;
 Matches 106; Conservative 0; Mismatches 100; Gaps 0;
 Indels 0;

Qy 88 ATTTGAGTCCAAATAATTATTTCATGCTATGAGAAAGCACATGACTAAAGTAA 147
 Db 3168 ATATAAACTCACCTTAACTTAAACCTAAACCTAAACCTAAACCTAAACCTAA 3109
 Qy 148 GCTTAATCCCTAAACCTAACATACAAAGGAGATGACATCCACATCCACATCTTAAATT 207
 Db 3108 AAACCCAAAAATTCAAAACCAACATAAATAAAACCCCCCACCTCTACTAAA 3049
 Qy 208 AGTCTTGGTAGAATTGAAACTTAACCTTAACTCTAACTTAAACCTAACTTAA 267
 Db 3048 ATACAAATAAAATAAACCTTAATTAACCTTAACTCTAACTTAAACCTAACTTAA 2989
 Qy 268 AAAATAAAAGAAAGTCTAACATGAA 293
 Db 2988 AAATAAAATACTAACCTAAACCTAA 2963

RESULT 5
 US-10-239-676-210/C
 ; Sequence 210, Application US/10239676
 ; Publication No. US20030082609A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
 ; FILE REFERENCE: 5013.1003
 ; CURRENT APPLICATION NUMBER: US/10/239,676
 ; PRIORITY APPLICATION NUMBER: PCT/EP01/03968
 ; DE 10019173.8
 ; DE 10032529.7
 ; DE 10043826.1
 ; PRIORITY FILING DATE: 2001-04-06
 ; LENGTH: 11812
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:

APPLICANT: SHIGENOBU, SHUJI
; APPLICANT: WATANABE, HIDEMI
; APPLICANT: HATORI, MASAHIRO
; APPLICANT: SAKAKI, YOSHITOKU
; TITLE OF INVENTION: GENOME DNA OF BACTERIAL Symbiont OF APHTIDS
; FILE REFERENCE: 081356/0159
; CURRENT APPLICATION NUMBER: US/09/790,988
; CURRENT FILING DATE: 2001-02-23
; PRIORITY APPLICATION NUMBER: JP2000-107160
; PRIORITY FILING DATE: 2000-04-07
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO: 1
; LENGTH: 640581
; TYPE: DNA
; ORGANISM: Buchnera sp.
; US-09-790-988-1

Query Match Score 39.4; DB 11; Length 640681;
Best Local Similarity 52.8%; Pred. No. 59; Indels 0; Gaps 0;
Matches 85; Conservative 0; Mismatches 76;

QY 190 ACAGAAAAAAATTCTAATTAATCTCTTGGGTTGAAACTGATACTACATAA 249
Db 611989 AAATATAATTGAAAATATTGAGAAAATATTATAATAAGTAA 612048

QY 250 TTACACATTGCAATAAATATAAGAAAGTCTAACATGAGACTAGTCTAACATG 309
Db 612049 TCATATGCTATTATAATAGATAATACTTATAAAGTATTGTTTATAGC 612108

QY 310 AAGACTAGTCACGAACTCTACCTTACATTCACAAAGCTT 350
Db 612109 AAGAATATAGAAACAAATATTATAAGAGCTT 612149

RESULT 14

US-10-239-676-41/C
; Sequence 41, Application US/10239676
; Publication No. US20030082609A1

GENERAL INFORMATION:

; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPERBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1000
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIORITY APPLICATION NUMBER: PCT/EP01/03968
; DE 10019173.8
; DE 10032529.7
; DE 10043826.1
; PRIORITY FILING DATE: 2001-04-06

DE 2000-04-06

2000-04-07

2000-06-30

2000-09-01

SEQ ID NO: 41

LENGTH: 6167

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)

US-10-239-676-41

NUMBER OF SEQ ID NOS: 228

SEARCHED:

QY 336 ATTICACAAAGGCTTAGACTTCCACAATGGAGATTATCCCATGGACTGATGGACACCA 395
Db 2003 TATTCACCAACTTAAACGAAACAAACGAAATCAAATCAAATCGAACCA 2004

QY 396 TCCAATTATCCCTATAATACCTGCCATTGCCCTCCAGCTCATCTAACCTAAA 455
Db 2003 TCCCTAACTAACAGATAAACCCATCTACTAAATAAAATTAACCGAACAGT 1944

QY 456 AGAACACACACACATAT 474
Db 1943 AATAATAACCACCTATAAT 1925

RESULT 15
US-09-239-842A-4290
; Sequence 4290, Application US/09338842A
; Patent No. US20020160378A1

GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI100-3

CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIORITY NUMBER: US 60/227,866
; PRIORITY FILING DATE: 2000-08-24
; PRIORITY NUMBER: US 60/264,647
; PRIORITY FILING DATE: 2001-01-16
; PRIORITY NUMBER: US 60/300,111
; PRIORITY FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO: 4290
; LENGTH: 864
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana

Query Match Score 38.8; DB 11; Length 864;
Best Local Similarity 64.4%; Pred. No. 3.6;
Matches 58; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 193 GAAAAATCTAACTTGTGAGAAATGAACTGATACCTCATATAATTAA 252
Db 84 GAAAGAGCTAAATAGCTTATGCTGTGAGGGTCAAAACTAAAGCTTT 143

QY 253 CAACTTTGCAATAAATATAAGAAAGT 282
Db 144 AGCTCTTAGTAAATAATGTGATAGT 173

Search completed: July 20, 2003, 06:21:49
Job time : 174 secs

Query Match Score 39; DB 15; Length 6167;
Best Local Similarity 49.7%; Pred. No. 8.3;
Matches 99; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

Qy 276 AGAAAGTCTAACATGAAGACTAGTCTAACATGAAGACTAGTCCAGAACACTGTTACCTT 335
Db 2123 ATAAAAATAAACCTAAATATAACCCACATAAAACACAAATAACTCATCTA 2064

GenCore version 5.1.6
 Copyright (c) 1993 - 2003 Compugen Ltd.
 M nucleic - nucleic search, using sw model
 run on: July 20, 2003, 04:50:04 ; Search time 1127 Seconds
 sequence: 1 gaattcttattgcgacatcgaa.....acaacacacaatcatg 475
 scoring table: IDENTITY_NGC
 Gapop 10.0 , Gapext 1.0
 searched: 1615406 seqs, 8097743376 residues
 total number of hits satisfying chosen parameters: 32308132
 post-processing: Minimum Match 0%
 Maximum Match 100%
 Listing first 45 summaries
 database : EST:
 1: em_estba:
 2: em_estnum:
 3: em_estln:
 4: em_estru:
 5: em_estcy:
 6: em_estpl:
 7: em_estro:
 8: em_htc:
 9: gb_est1:
 10: gb_estL:
 11: gb_htc:
 12: gb_estC:
 13: gb_estE:
 14: gb_estS:
 15: em_estfun:
 16: em_escom:
 17: gb_gss:
 18: em_gss_hum:
 19: em_gss_inv:
 20: em_gss_Pln:
 21: em_gss_vrt:
 22: em_gss_fun:
 23: em_gss_mam:
 24: em_gss_mus:
 25: em_gss_other:
 26: em_gss_pro:
 27: em_gss_rnd:
 Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed,
 > and is derived by analysis of the total score distribution.

Result No.	Score	Query	Match	Length	DB	ID	Description
1	61.2		12.9	1101	17	CNS0100X	AI098319 Drosophila
2	59		12.4	1101	17	CNS039G	AL063921 Drosophila
3	48.4		10.0	1101	17	CNS1180R	AL106391 Drosophila
4	47.4		10.0	1101	17	CNS039G	AL063925 Drosophila
5	45.6		9.6	1043	17	CNS0145P	AL103735 Drosophila
6	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
7	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
8	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
9	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
10	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
11	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
12	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
13	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
14	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
15	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
16	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
17	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
18	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
19	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
20	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
21	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
22	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
23	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
24	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
25	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
26	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila
27	45.5		9.6	1043	17	CNS0145P	AL103735 Drosophila

SUMMARIES

No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

REFERENCES

result		Score	Query	Length	DB	ID	Description
No.	No.						
3	1	61.2	12.9	1101	17	CNS0100X	AL098379 Drosophil
3	2	.59	12.4	1101	17	CNS0039G	AL063921 Drosophil
3	3	48.4	10.2	1101	17	CNS0100R	AL0108741 Drosophil
3	4	47.4	10.0	1101	17	CNS0039G	AL063921 Drosophil
3	5	45.6	9.6	1043	17	CNS0115P	AL103735 Drosophil
3	6	45.0	9.6	1043	17	CNS0115P	AL103735 Drosophil

ALIGNMENTS

RESULT 1	CNS0100X/C	CNS0100X	1101 bp	DNA	linear	GSS 26-JUL-1999
LOCUS	Drosophila melanogaster	Genome survey sequence	SP6 end of BAC			
DEFINITION	BACN03G04 of DrosBAC library from Drosophila melanogaster (fruit fly)	genomic survey sequence.				
ACCESSION	AL098379					
VERSION	AL098379.1	GI:5609990				
KEYWORDS	GSS.					
SOURCE	Drosophila melanogaster.					
ORGANISM	Drosophila melanogaster					
EUKARYOT	Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neopiera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila.					
1	(bases 1 to 1101)					
AUTHORS	Genoscope.					
TITLE	Direct Submission					
JOURNAL	Submitted (23-JUL-1999) Genoscope - Centre National de Séquençage					
JOURNAL	BP 191 91006 EVRY cedex FRANCE (E-mail : seqfr@genoscope.cnrs.fr)					
COMMENT	- Web : www.genoscope.cnrs.fr					
	Determination of this BAC-end sequence was carried out as part of a collaboration with the European Drosophila Genome Project (EDGP) - http://www.edgp.ebi.ac.uk/ . This Drosophila melanogaster BAC library (Dros BAC) was made by Alain Billaud at CEPH (Centre d'Etude du Polymorphisme Humain) with funding provided by a MRC project grant. The DNA was prepared from embryos by Alain Bucheton and Geneviève Payan. It has been constructed in the vector					

ESTIMATE	4	REFINEMENT	CNSNS039G/C	DEFINITION	CNS039G Drosophila melanogaster genome survey sequence TET3 end of BAC # BACR010 of RCT-98 library from Drosophila melanogaster (fruit fly). Genomic Survey Sequence.	SEQUENCE	GSS 03-JUN-1999
ACCESSION	AL063911	VERSION	1	SOURCE	GR:4941778	AUTHORS	
VERSIONS	AL063911.1	ORGANISM	Drosophila melanogaster.	REFERENCE		JOURNAL	
REMARKS		ORGANISM	Drosophila melanogaster. Bukaryota; Metacoda; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidae; Drosophilidae; Drosophila. 1 (bases 1 to 1101)	COMMENT			
GENOTYPE		DIRECT SUBMISSION	Submitted (02-JUN-1999) Genoscope - Centre National de Sequençage : BP 191 91006 EYVY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)	INDIRECT SUBMISSION			
WEB		WEB	http://www.genoscope.cns.fr	COMMENT			
		DETERMINATION	Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP).	COMMENT			
		RESULTS	The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammone in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-88 and was constructed by partial BAC/RI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp. The same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the BACPAC Resource Center, or filters for hybridization from the BACPAC Resource Center can be	COMMENT			

Qy	157 CCTAAACTCAATAAACAGGATGACATCCAGAAAATTCTTAATTAGTGTTCGC 216	Qy	212 TTGCGGTGTTAGAATTGGAACTGAATACCTACATTATTACACTTGTGCAAATAAAAT 271
Db	784 TAAAAAAAATTWAAACTTAAATAAAAARAAAAAATTTTAHCCPATTAA 843	Db	288 AAAAAGAAAAAAATTAAATAGAGAAAATCCATTATAAAATTAAAATAAAAT 347
Qy	217 GTGTAGAAATTGGAAACTGAATACCTACATTAAATTACAACCTTGTGCAAATAAAAT 276	Qy	272 ATAAAGAA 279
Db	844 ATWAAAATTTGGWTTTAAAAAATAATTTHTAATTTAAATAAAATAAAATAAA 903	Db	348 ATTAAAA 355
Qy	277 GAAAGCTCTACATGAGACTAGTCTAACATGAGAATGTCGACCTTA 336		RESULT 9
Db	904 WWWWWWWAAAWAAAHCCCCAAAMCTTAAAAA - AACCT 961	AL513901	AL513901 LTI_NFL006_PL2 Homo sapiens mRNA linear EST 13-FEB-2001
Qy	337 TTCCACAAAGGCTTAGCTTCCACAAATGAGATTATCCGACTGGACACAT 396	LOCUS	AL513901 mRNA sequence.
Db	962 TWAAAATAAAAYWHAMAAAAMMMMAAMCCTYYAAAAMACTWAMAAAAMA 1021	DEFINITION	AL513901 AL513901.1 GI:12777395
Qy	397 CCAAATATATCCCTTAAATAACTGCCCATCCCTCCAGACTCATTAACCTAA 456	VERSION	EST.
Db	1022 ATAHHAMMCMCTAWAAAAMMMCCAAAMMAACAWNCATNMAMCMMCMMMC 1081	KEYWORDS	
Qy	457 C 457	SOURCE	Homo sapiens
Db	1082 C 1082	ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Buthidae; Primates; Catarrhini; Hominidae; Homo.
		REFERENCE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
		AUTHORS	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
		TITLE	Full-length cDNA libraries and normalization
		JOURNAL	Unpublished (2001)
		COMMENT	Contact: Genoscope Genoscope - Centre National de Séquençage BP 191 91006 Evry cedex - France Email: sequenc@genoscope.cnrs.fr, Web : www.genoscope.cnrs.fr.
		FEATURES	Location/Qualifiers
		SOURCE	1. bases 1 to 638
			/organism="Homo sapiens"
			/db_xref="taxon:9606"
			/clone_lib="LTI_NFL006_PL2"
			/tissue_type="placenta"
			/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dh) primer. Five prime end enriched, double-stranded DNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Fang Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1).301 610 8371 Email: fliang@lifetech.com URL : http://fulllength.invitrogen.com"
		BASE COUNT	347 a 38 c 15 g 141 t 77 others
		ORIGIN	BP 191 91006 Evry cedex - France
		Query Match	9.3%; Score 44; DB 9; Length 638;
		Best Local Similarity	39.5%; Pred. No. 2,3; Mismatches 119; Indels 0; Gaps 0;
		Matches	98; Conservative 31; Mismatches 119; Indels 0; Gaps 0;
		Qy	119 ATGGAGAGGCACTGACTTAAGTATAGCTTAATCCCTAAACTTAATCAACAGG 178
		Db	174 ATAAMMAAAAAMAAAAMAAAATATTACMAAAAAMAAAAMAAAAM 233
		Qy	179 ATGGCACATCCACGAAATCTAAATGCTTGGTGTAGAAATTGGAAACTGAT 238
		Db	234 GGGAAAWTTWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAAWAA 293
		Query Match	9.3%; Score 44; DB 9; Length 358;
		Best Local Similarity	39.5%; Pred. No. 2,3; Mismatches 90; Indels 0; Gaps 0;
		Matches	98; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
		Qy	239 ACCATACATTAATTACAACCTTGTGAAATAAAGATAAGAAACTCTACATGAGCTA 298
		Db	294 AATPAAAATTTAGAACATGTTCAATGTGTTGACAAATTATTGTC 353
		Qy	299 GTCTAACATGAAAGACTAGTCCACGAAACTCTGTTTACATGAGCTTC 358
		Db	354 AAAAAGAAAAAMMMMAACCCCCCTTTAMCCAMCCMATAAHMCHIWCCCCCA 413
		Qy	359 CACAAATC 366
		Db	414 AAAAAC 421

RESULT 10	CNS011EU/C	Drosophila melanogaster	860 bp	DNA linear	GSS 26-JUL-1999	BACN37B03 of DrosBAC library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.
DEFINITION	BACN06H18 of DrosBAC library from <i>Drosophila melanogaster</i> (fruit fly), genomic survey sequence.					
ACCESSION	AL100176					
VERSION	AL100176.1					
KEYWORDS	Drosophila melanogaster					
SOURCE	Drosophila melanogaster					
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydriidea; Drosophilidae; Drosophila.					
REFERENCE	GSS.					
AUTHORS						
TITLE						
JOURNAL						
COMMENT						
FEATURES						
source						
base count	202 a	202 a	106 c	116 g	306 t	130 others
origin						
Query Match	9.2%	Score 43.8;	DB 17;	Length 1101;		
Best Local Similarity	41.2%	Pred. No. 2.6;	Mismatches 14;	Indels 0;	Gaps 0;	
Matches	131;	Conservative				
QY	54	TTCTGTGCGACAAATACTGGTCCTGGTGAAGTCCAAATAATTTC	113	Db	69 ACATACTGTCCTGCCTCAATAATTTCATGTCATGAGAAG 128	
Db	319	TNNWTGCNCGGWTTCTATRWCGGTTNNNATATGGCGT	--AAAATAGNTGAA 263	Db	299 ACAGATTAGCTGAGTCACANTAAAAATAAGAAATAATAAAATAAA 240.	
QY	114	ATGTCATGAGAGGCCACATGACTAAAGTAGCTTAATTCCTAAACTCAATACAA 173		QY	129 CACATGACTAAAGTAATTAGCTTAATTCCTAAACTCAATAAAAGGAGATCACATC 188	
Db	262	ATGAAAATAATGCAAMATRCAAAGAACTCCCTCATWCMAGAWCTTC 203		Db	239 MACAAAAGAAAAAAATAATAATAATAATAATAATAATAATAAAAWAA 180	
QY	174	ACGGAGTACACATCCACAGAAAAATTCCTAAATTCTGTTGCGTGAAATTGAAAC 233		QY	189 CACGAAAAAATCTTAATTAGTCCTTGGCTGTTGAAATTGAAACTCTACATTA 248	
Db	202	WWMAAAATAATAAGAAATAAAATAACTASATTTCAAAATKCPAAWWTAAA 143		Db	179 AWAAAAMAAAAAAAGRAAAAMAAAATAAAATAAAATAAAATAAAATAAA 120	
QY	234	TGAAATACCTACATTAATTACAACCTTTCGAAATAAAATAAAAGAAATCTAACATGAA 293		QY	249 ATTACAACTTGGAAATAAAATAAAAGAACTGCTAACATGAAACTAGTCCTACAT 308	
Db	142	AMAAATAAAATAAAATAAAAWWWKAATAATTATAAAATAAAATAAAATAAA 83		Db	119 AAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAAATAAA 60	
QY	294	GACTAGTCTAACATGAAAGACTAGTCACGAACTCTGACCTTATICACACAAAGGCTTGA 353		QY	309 GAAGACTAGTCACGAACTCTGCTACCTTATT 338	
Db	82	AGATAGTTATAATTAACTTCCTTGTGTTAKAMTTIAWHAWTTATAATWTTAAGT 23		Db	59 CATCATCTNTCTCATGTTATCTGTTTTATT 30	
QY	354	CTTCACAAATGAGAT 371				
Db	22	CATTAAATTATTATGAT 5				
RESULT 11	CNS017PB/C	Drosophila melanogaster	1101 bp	DNA linear	GSS 26-JUL-1999	
DEFINITION	CNS017PB					
LOCUS						
DEFINITION	Drosophila melanogaster genome survey sequence SP6 end of BAC					
ACCESSION	AL0594400					
VERSION	AL0594400.1					
KEYWORDS	GSS.					
SOURCE						
ORGANISM	Drosophila melanogaster					
DEFINITION	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;					

